

**Análisis predictivo del impacto de proyectos de desarrollo vial en precios de viviendas:
RegioTram de occidente**

Daniel Felipe Triana Sanguña

Yenni Paola Reyes Tao

Asesor

Felipe Alexander Pipicano Guzmán

Universidad Nacional Abierta y a Distancia UNAD
Escuela de Ciencias Básicas, Tecnología e Ingeniería CBT
Ciencia de Datos y Analítica

2024

Resumen

La dinámica inmobiliaria en Cundinamarca, especialmente en municipios cercanos a Bogotá como Mosquera, Madrid, Funza y Facatativá, ha crecido notablemente. Este auge está impulsado en parte por la construcción del Regiotram de Occidente, que podría influir en los precios de las propiedades, generando ciclos de auge y posibles burbujas. Por ello, es esencial estudiar las variables que afectan los precios de las viviendas en esta región, como la ubicación, el estrato socioeconómico y la infraestructura de transporte. Los modelos de machine learning son herramientas útiles para predecir estos precios.

El objetivo de la monografía es analizar los modelos que se han utilizado para estudiar el impacto de proyectos férreos en el mercado inmobiliario. Se revisarán estudios previos y se identificarán las variables clave que influyen en la relación entre infraestructura vial y precios de las viviendas. Además, se espera tener una idea de la posible variación en los precios de las propiedades cercanas al Regiotram de Occidente por medio de literatura basada en estudios de modelos de machine learning aplicados en otras ciudades. También se discutirán los beneficios y retos de estos proyectos, como la mejora en la conectividad, la reducción de tiempos de desplazamiento y los posibles impactos negativos en la planificación urbana.

Palabras clave: Regiotram, Dinámica Inmobiliaria, Precios de Las Viviendas, Machine Learning, Infraestructura de Transporte.

Abstract

The real estate dynamics in Cundinamarca, particularly in small towns near Bogotá such as Mosquera, Madrid, Funza, and Facatativá, have grown significantly. This boom is partly driven by the construction of the Regiotram de Occidente, which could influence property prices, generating boom cycles and potential bubbles. Therefore, it is essential to study the variables that affect housing prices in this region, such as location, socioeconomic status, and transportation infrastructure. Machine learning models are useful tools for predicting these prices.

The objective of the monograph is to analyze the models that have been used to study the impact of rail projects on the real estate market. Previous studies will be reviewed, and key variables influencing the relationship between road infrastructure and housing prices will be identified. Additionally, there is an expectation to gain insight into the potential variation in property prices near the Regiotram de Occidente through literature based on machine learning models applied in other cities. The benefits and challenges of these projects will also be discussed, including improvements in connectivity, reductions in travel times, and potential negative impacts on urban planning.

Keywords: Regiotram, Real Estate Dynamics, Housing Prices, Machine Learning, Transportation Infrastructure.

Tabla de Contenido

Introducción	8
Justificación	10
Objetivos.....	12
Objetivo General.....	12
Objetivos Específicos.....	12
Planteamiento del Problema	13
Marco Conceptual y Teórico	15
Dinámica Inmobiliaria en la Sabana de Occidente	15
Influencia de la Infraestructura Vial en los Precios de las Viviendas.....	17
Modelos de Machine Learning en el Análisis Inmobiliario.....	20
Modelos de Regresión	20
Modelos Avanzados.....	21
Metodologías de Optimización y Evaluación.....	21
Análisis Complementarios	21
Burbujas Inmobiliarias: Causas y Efectos	22
Contexto del Proyecto RegioTram de Occidente.....	24
Descripción del Proyecto	24
Impacto Esperado en los Municipios Cercanos a Bogotá.....	26
Situación Actual del Mercado Inmobiliario en Mosquera, Madrid, Funza y Facatativá.....	26
Modelos de Machine Learning para Analizar el Impacto del Regiotram de Occidente en la	
Valorización Inmobiliaria en sus Áreas Cercanas	34
Diseño del Estudio	35

Recolección de Datos.....	37
VARIABLES ANALIZADAS	40
Aplicación de Modelos de Machine Learning	44
Ventajas y Desventajas de los Modelos.....	44
Regresión Lineal Múltiple	44
Algoritmo Híbrido	45
XGBoost	46
Comparación con Estudios Previos	47
Kuala Lumpur	47
Delhi.....	47
Conclusión y Modelo Recomendado	48
Resultados	49
Predicciones de Variación en los Precios de las Viviendas	49
Identificación de las Variables Clave: Transporte y Precios Inmobiliarios.....	50
Comparación de Resultados con Estudios Previos	50
Discusión.....	52
Impacto del Regiotram de Occidente en el Mercado Inmobiliario.....	52
Beneficios y Retos de la Infraestructura Vial en Cundinamarca	52
Efectos sobre la Conectividad y Reducción de Tiempos de Desplazamiento	53
Riesgos de Burbujas Inmobiliarias y Planificación Urbana Sostenible.....	54
Conclusiones	55
Recomendaciones	58
Referencias Bibliográficas	60

Lista de Tablas

Tabla 1 Precios m^2 Integral y Variaciones Porcentuales	16
Tabla 2 Número de Personas por Hogar	28
Tabla 3 Nivel de Ingresos por Hogar \$0 – \$3,5M.....	28
Tabla 4 Nivel de Ingresos por Hogar \$3,5M- Más de \$9M	29

Lista de Figuras

<i>Figura 1</i> Mapa de Trazado de RegioTram de Occidente.....	25
<i>Figura 2</i> Población de la Sabana de Occidente.....	27
<i>Figura 3</i> Condiciones de Pobreza.....	29
<i>Figura 4</i> Saldos Migratorios -2021.....	30
<i>Figura 5</i> Vivienda.....	31
<i>Figura 6</i> Crecimiento de la Huella Urbana.....	32

Introducción

La dinámica inmobiliaria en la región de Cundinamarca ha experimentado un auge significativo en los municipios cercanos a Bogotá, destacando Mosquera, Madrid, Funza y Facatativá. La construcción del Regiotram de Occidente ha influido en este aumento, pero también puede generar ciclos de auge y burbujas de precios. Es relevante explorar el impacto a largo plazo en la economía local y la vida de las personas en la Sabana. Esto genera que en la toma de decisiones de inversión en vivienda sea fundamental, estudiar las variables que inciden en la determinación de precios de las viviendas en la sabana de occidente. Los modelos de machine learning pueden ayudar a determinar el precio de las viviendas, considerando variables como la ubicación, el estrato socioeconómico y la infraestructura de transporte.

El objetivo general de la monografía es analizar los diversos modelos utilizados en la literatura determinando las variables y los efectos de la construcción de líneas de trenes en los precios de las viviendas. Los objetivos específicos son identificar las tendencias de Machine Learning, revisar estudios previos, catalogar las variables que más influyen en la relación entre la construcción de proyectos viales y los precios de viviendas.

El marco conceptual y teórico de la monografía se centra en la infraestructura de transporte y su impacto en los precios de las viviendas. Los proyectos férreos, como el Regiotram de occidente, pueden afectar los precios de las propiedades residenciales y comerciales en áreas aledañas. Es importante considerar la disminución de tiempos de transporte, facilidad de acceso a comercios, sitios de interés y ciudades desarrolladas, así como el bajo impacto ambiental de algunos proyectos de infraestructura vial.

Se espera analizar los modelos de Machine Learning más utilizados en estudios previos en diferentes ciudades del mundo, la variación que podrían tener los precios de las propiedades

residenciales y comerciales en áreas aledañas de no más de 1.5 kilómetros de distancia de la estación de tren por donde va a pasar el Regiotram de occidente. Esto permitirá a los inversores y compradores de viviendas tomar decisiones informadas y fundamentadas en la relación entre la construcción de proyectos viales y los precios de las viviendas.

Además, se analizarán los beneficios y desafíos que implica la construcción de proyectos de infraestructura vial en la región, como el aumento de la conectividad y la mejora de la calidad de vida acortando los tiempos de desplazamiento de las personas que laboran en Bogotá y viven en estas ciudades aledañas. También se discutirán los posibles impactos negativos, como la planificación urbana sostenible y la participación ciudadana en la toma de decisiones.

Justificación

Invertir o cambiar de residencia es una decisión determinante en toda familia, por lo general las familias buscan tener vivienda propia basados en condiciones climáticas, ambientales tanto internas como externas y de cercanías con las capitales, sobre todo en cuestiones de vías de transporte para el desplazamiento, es por esto, que tener claridad sobre la tendencia en los precios de las viviendas de algún sector en el que se quiera vivir es fundamental, sobre todo si los precios establecidos por catastro no son acordes a los valores comerciales de la zona (Grajales, 2019).

Estudiar las variables que inciden en la determinación de precios de las viviendas en la sabana de occidente es de gran importancia para mejorar la toma de decisiones de inversión, además de las condiciones de dichas viviendas permitiría negociar de manera justa y equitativa la compra y venta de dichos inmuebles con un plus que juega un papel relevante y es que el tema de la valorización por la construcción de un proyecto vial como el Regiotram de occidente, podría determinar el alza en los precios de las viviendas (Rodríguez et al., 2021).

Tradicionalmente los modelos de machine learning, tales como regresiones lineales y logísticas para la determinación de los precios de las viviendas se ha enfocado en variables como el tipo de casa, la ubicación, el estrato socioeconómico, los años de construido, el área, el número de baños, el número de habitaciones, la antigüedad de la vivienda y si cuenta con parqueadero propio o comunal, arrojando resultados con rendimientos óptimos que ayudan a determinar el precio (Selim, 2009). Sin embargo, Puentes y Garza, (2023), destacan que si bien es importante tener en cuenta las características intrínsecas de las viviendas, la mayoría de las contribuciones de la literatura académica identifican un efecto positivo en el precio de las viviendas a causa de la infraestructura de transporte y que es determinante evaluar la hipótesis

acerca de los efectos tanto en el corto como en el largo plazo, asegurando que ante la incertidumbre que tal vez puede ocasionar la volatilidad de los precios en el corto plazo, al final dichos proyectos son determinantes para la valorización de los predios.

Teniendo en cuenta esto, este trabajo busca analizar y revisar los diversos modelos utilizados en la literatura para determinar las variables y los efectos que tendría la construcción de líneas de trenes en los precios de las viviendas de la Sabana de occidente, específicamente en los municipios de Mosquera, Madrid, Funza y Facatativá, ante la implantación de un sistema de transporte público regional para dichos municipios que conectará la sabana al centro de Bogotá, reduciendo los tiempos de desplazamiento e incrementando las posibilidades de transporte de sus habitantes y cómo este impactará y hasta qué punto lo hará basados en estudios de aplicación de algoritmos de machine learning para la determinación de los precios de las viviendas con influencia de desarrollos viales.

Objetivos

Objetivo General

Analizar los diversos modelos utilizados en la literatura para determinar las variables que influyen y los efectos de la construcción de líneas de trenes como el RegioTram en los precios de las viviendas en la sabana de occidente.

Objetivos Específicos

Identificar las tendencias de Machine Learning y variables que determinan el precio de las viviendas basados en referentes bibliográficos y datos históricos.

Comparar los estudios previos que han investigado el impacto de la construcción de líneas de trenes en los precios de las viviendas en diferentes contextos geográficos y socioeconómicos.

Catalogar las variables que más influyen en la relación entre la construcción de proyectos viales y los precios de viviendas.

Planteamiento del Problema

Durante los últimos años se ha visto cómo la dinámica inmobiliaria en la región de Cundinamarca ha tomado auge en los municipios adyacentes a Bogotá, dada la cercanía. Dentro de lo que Rodríguez et al., (2021), ha denominado el primer anillo geográfico de Bogotá, de los seis municipios que aportaron la mayor cantidad de oferta de viviendas, se encuentran en su predominancia el municipio de Mosquera, Madrid, Funza y Facatativá, siendo el primero con alrededor de 4.500 viviendas en ofertas durante 2017 a 2020, con una variación positiva de 7% respecto a 2018, con un incremento entre el 2.2% y 7.8% en el metro cuadrado, lo cual ha impactado que la valorización de los predios esté incrementando, presentándose un fenómeno especulativo en los precios (Grajales, 2019).

Lo anterior, sumado a la construcción de importantes proyectos viales como el Regiotram de Occidente, como corredor férreo que permitirá un proceso de revitalización que busca resolver las necesidades de movilidad de la población creciente en los municipios de la sabana de occidente (Alonzo, 2017), ha hecho que sea de nuestro interés durante el proceso de investigación de esta propuesta, entender si el incremento en los precios de las viviendas en la Sabana de occidente está afectado por el nuevo tren de cercanías.

El problema específico que se ha identificado es la valorización (alza en los precios) de las propiedades en la zona de la Sabana, específicamente en los municipios de Mosquera, Funza, Madrid y Facatativá, debido a la promesa de un tren de cercanías, incluso antes de que se construya. Este fenómeno puede ser impulsado por las constructoras que ofrecen beneficios relacionados con la facilidad de transporte, al analizar el impacto del sistema de tren ligero en las viviendas en Charlotte, Carolina del Norte, desde 1997 a 2008, donde la proximidad a una estación hizo contribuir modestamente a la variación del precio de la vivienda (Alonzo, 2017).

Lederman, (2023), por su parte, asegura que, al tener implicaciones significativas para los compradores de vivienda y el mercado inmobiliario en general, se podría tener ciclos de auge, si los precios de las propiedades aumentan basándose solo en la promesa de un tren de cercanías, esto podría resultar en una especulación inmobiliaria y posiblemente en una burbuja de precios. Además, podría excluir a los compradores potenciales que ya no pueden permitirse comprar en la zona debido al aumento de los precios. Tal y como lo expone Abido et al., (2021), la construcción de infraestructuras de transporte puede tener un impacto en las propiedades cercanas, la economía y el medio ambiente. Específicamente, se han realizado estudios para entender cómo el valor de las propiedades puede cambiar debido a un mejor acceso al transporte.

Las deficiencias que se han identificado respecto al alza en los precios de las viviendas producto de la promesa de construcción de un tren de cercanías, está relacionado con varios factores que involucran el desplazamiento de residentes de bajos ingresos, que no cuentan con el suficiente poder adquisitivo para la adquisición de una vivienda, tal vez las especulación inmobiliaria, el incremento en el costo de vida, la exclusión socioeconómica y la presión existente sobre la infraestructura, que son temas tanto sociales como económicos que no están contemplados en un modelo cuando se trata de estimar el precio (Alonzo, 2017).

Algunos aspectos del problema que podrían no haberse explorado completamente incluyen el impacto a largo plazo de este tipo de valorización de propiedades en la economía local y la vida de las personas en la zona de la Sabana (Escobar, 2019). Además, podría ser útil investigar cómo se comparan estas tendencias con otras regiones que han experimentado proyectos de desarrollo similares. Mediante una monografía podríamos ayudar a resolver ¿Qué variables inciden en la determinación del precio de la vivienda en la sabana de occidente y tiene alguna afectación la construcción del Regiotram?

Marco Conceptual y Teórico

Dinámica Inmobiliaria en la Sabana de Occidente

Desde que las civilizaciones empezaron a desarrollar sus proyectos de infraestructura, las áreas aledañas a estos se han visto afectadas de diferentes maneras. Parques, bibliotecas, escenarios deportivos, centros comerciales, estaciones de transporte y terminales, son algunos de los desarrollos de infraestructura más comunes en ciudades y regiones. Dentro de estos se encuentran los proyectos de infraestructura vial y de transporte, que tal y como lo indica González et al., (2015) potencian el desarrollo regional por medio de un proceso de cambio estructural que está directamente relacionado con la disponibilidad de infraestructura de transporte y comunicaciones, por eso en la medida en que las relaciones espaciales se consoliden, las economías de escala se potencian. En los proyectos de desarrollo vial o de transporte, encontramos el carretero, marítimo, fluvial y férreo (Gil et al., 2022). Los proyectos férreos, objeto de este estudio, están relacionados con transporte de personas, animales, mercancías de un lugar a otro por medio de vías férreas, metros y comunicación local (Gil et al., 2022).

De acuerdo a lo que establece García, (2021), la oferta de vivienda en la región de Cundinamarca, específicamente en la sabana de occidente, que para efectos de este estudio contempla los municipios de Madrid, Funza, Mosquera y Facatativá, la oferta inmobiliaria es significativa, dada la cercanía al polo de desarrollo económico que constituye la mayor fuente de empleo de la región por estar cercano a la capital, condiciones de ambiente favorables, seguridad, precios y área de las viviendas, entre otros, son propicias para impulsar el mercado inmobiliario en estas zonas del país, de ahí que los cuatro municipios del occidente de la ciudad, aportaron durante los años 2017 a 2020 la mayor cantidad de ofertas de vivienda, con predominancia del municipio de Mosquera, con alrededor de 4.500 viviendas, con una variación

positiva del 7% en 2018 respecto al 2017, no obstante si bien todos los municipios tuvieron un decrecimiento en las ofertas del año 2020 por efectos de la pandemia, también es importante destacar que su disminución fue relativamente baja, si lo comparamos con otros municipios de Cundinamarca. Así mismo, a partir del tipo de vivienda de los inmuebles ofertados, la cual se relaciona con la cantidad de apartamentos y de casas, permitió determinar que, a lo largo de los cuatro años de estudio en estos municipios, predomina la oferta de apartamentos, alcanzando cifras desde 300 hasta 1.300 ofertas anuales, concentrado en su gran mayoría en los municipios de Mosquera y Funza.

En cuantos a los precios de las viviendas por m², tal como se muestra en la Tabla 1, el estudio encontró que Facatativá registró los precios más bajos del mercado en \$1.279.070 y \$1.380.000 para los años 2020 y 2019, respectivamente, mientras que en los demás municipios se mantuvo casi que estable. Sin embargo, llama la atención el municipio de Mosquera, donde el m² es más elevado, pero aun así sigue siendo más bajo que en la ciudad de Bogotá donde el m² oscila entre \$2.673.077 y \$6.326.531, dependiendo el estrato y la localidad (Rodríguez et al., 2021).

Tabla 1

Precios m² Integral y Variaciones Porcentuales

Año	Valor, m ²	Mosquera	Madrid	Funza	Facatativá
2017	Mínimo	\$ 1.810.345	\$ 1.716.418	\$ 1.764.706	\$ 1.617.647
	Mediana	\$ 2.304.348	\$ 2.361.111	\$ 2.406.996	\$ 2.227.183
	Máximo	\$ 3.658.537	\$ 3.444.444	\$ 3.245.614	\$ 2.857.143
2018	Mínimo	\$ 1.810.345	\$ 1.714.286	\$ 1.779.661	\$ 1.617.647
	Mediana	\$ 2.466.667	\$ 2.409.241	\$ 2.410.714	\$ 2.169.811
	Máximo	\$ 3.630.435	\$ 3.563.218	\$ 3.285.714	\$ 3.060.000
2019	Mínimo	\$ 1.940.299	\$ 1.515.152	\$ 1.833.333	\$ 1.380.000

	Mediana	\$ 2.713.075	\$ 2.500.000	\$ 2.530.120	\$ 2.241.379
	Máximo	\$ 3.818.182	\$ 3.555.556	\$ 3.314.607	\$ 3.076.923
	Mínimo	\$ 1.944.444	\$ 1.796.875	\$ 1.850.746	\$ 1.279.070
2020	Mediana	\$ 2.773.585	\$ 2.549.020	\$ 2.602.083	\$ 2.441.734
	Máximo	\$ 3.823.529	\$ 3.511.905	\$ 3.787.879	\$ 3.470.588
	Var. Min	0,00%	-0,10%	0,80%	0,00%
2017-2018	Var. Med	7,00%	2,00%	0,20%	-2,60%
	Var. Máx	-0,30%	3,50%	1,20%	7,10%
	Var. Min	7,20%	-11,60%	3,00%	-14,70%
2018-2019	Var. Med	10,00%	3,80%	5,00%	3,30%
	Var. Máx	4,70%	-0,20%	0,90%	0,60%
	Var. Min	0,20%	18,60%	0,90%	-7,30%
2019-2020	Var. Med	2,20%	2,00%	2,80%	8,90%
	Var. Máx	0,10%	-1,20%	14,30%	12,80%

Nota. Precios de las viviendas por m² en los municipios de la sabana de occidente de 2017 a 2020. Tomado de. Dinámica Inmobiliaria Bogotá - Región 2017 – 2020, por García, W, (2021).

Influencia de la Infraestructura Vial en los Precios de las Viviendas

Unos de los actores más afectados con respecto a los proyectos de desarrollo vial y de transporte en una región son los bienes raíces. Por lo general los precios de las propiedades son susceptibles a cualquier modificación, según Ayuso y Restoy, (2006), desde finales de la década de 1990, muchas economías industrializadas han experimentado un período prolongado de aumentos significativos en los precios inmobiliarios. En algunas de estas economías el proceso se ha vuelto particularmente agudo después de la fuerte corrección de los precios del mercado de valores iniciada a finales de 2000. Ahora se está volviendo una opinión generalizada que tal aumento en los precios inmobiliarios ha contribuido sustancialmente a la resiliencia de las

empresas privadas. Diferentes factores influyen a su variación de precios tales como, ubicación, facilidad de transporte, comercio y demás.

Para determinar la variación de estos precios en función específicamente de desarrollos de proyectos de infraestructura vial, nos apoyamos de una rama de la inteligencia artificial denominada Machine Learning la cual como afirma Abidoje et al., (2021), el aprendizaje automático es un paradigma que puede referirse al aprendizaje a partir de experiencias pasadas (que en este caso son datos previos) para mejorar el rendimiento futuro. El enfoque principal de este campo son los métodos de aprendizaje automático. El aprendizaje se refiere a la modificación o mejora de un algoritmo basado en “experiencias” pasadas de forma automática sin ninguna asistencia externa de un humano.

Autores como Debrezion et al., (2007), sostienen que las estaciones de ferrocarril, que actúan como nodos en las redes de transporte y lugares en un entorno urbano, tienen impactos ambientales y de accesibilidad que contribuyen al valor de la propiedad. Los estudios sobre los efectos de las estaciones de ferrocarril en el valor de la propiedad han arrojado resultados mixtos, variando desde un impacto negativo hasta uno positivo o insignificante. Igualmente, según Tyagi y Singhal, (2018), un resumen de estudios seleccionados sobre el tránsito ferroviario urbano y su impacto en las propiedades residenciales y comerciales ubicadas en un rango de 0 a 1000 m de distancia de los rieles del tren en varias ciudades metropolitanas alrededor del mundo mediante un modelo de regresión lineal llamado comparabilidad de la espacialidad, se descubrió que durante la fase de planificación y construcción, la proximidad a una estación de metro mostró una tendencia negativa en los precios de las propiedades comerciales. Sin embargo, durante la fase de operación, se observó una prima de precio significativa asociada con las propiedades

comerciales debido a la mejora de la accesibilidad. El tiempo, dividido en dos períodos: 2000-2004 (planificación y construcción) y 2005-2008 (operación).

Por otro lado, en países como Estados Unidos también encontramos documentación que soporta la idea de la variación de precios de acuerdo con la influencia de proyectos de infraestructura vial y de transporte. De acuerdo con Seo et al., (2019), mediante un estudio con el modelo de precios hedónicos llevado a cabo en Phoenix, Arizona sobre el impacto de las autopistas de alta velocidad y los trenes ligeros en valor de las propiedades comerciales, notaron que los valores de las propiedades comerciales están asociados positiva y significativamente con los beneficios de accesibilidad de los nodos de transporte. Los coeficientes de banda de distancia forman una curva típica de decaimiento de la distancia para ambos modos sin un efecto de donut de di amenidad detectable inmediatamente alrededor de los nodos. Analizando más a fondo el tema de estudio, los precios de la tierra están influenciados por factores como la integración, la conectividad y el ancho de la calle. Además, las zonas comerciales con buena integración y calles más anchas tienden a tener precios más altos, mientras que las calles locales estrechas generalmente tienen precios más bajos. Por último, el estudio ha construido con éxito un modelo para predecir el precio de la tierra urbana (Abdulla et al., 2023).

A pesar de que en gran parte de los estudios revisados se encuentra un impacto positivo en la variación de los precios de bienes raíces con respecto a los proyectos de infraestructura vial, otros autores como Puentes y Garza, (2023), dicen que la mayoría de las contribuciones en la literatura académica identifican un efecto positivo de la infraestructura de transporte en los precios de la tierra. Sin embargo, la dinámica a corto plazo de estos precios no ha sido analizada rutinariamente. Ellos encuentran ajustes a corto plazo volátiles que contradicen las predicciones neoclásicas, mientras que se asemejan a los ajustes de precios de la tierra espacial expuestos a la

incertidumbre radical. En esta tradición, los choques a corto plazo tienen efectos a largo plazo en la distribución espacial de los valores de la tierra debido a la incertidumbre radical.

Lo mismo descubrió Forouhar, (2016) en su estudio de estimación del impacto de las estaciones de metro en el valor de las propiedades residenciales en Teherán. Este evidenció un efecto negativo en vecindarios del norte, las estaciones de metro tienen un efecto general negativo en el valor de venta de las propiedades residenciales ubicadas en los vecindarios más acomodados del norte de Teherán. Mientras que tiene un efecto positivo en vecindarios del sur: Por otro lado, el efecto es positivo y alto para las propiedades ubicadas cerca de las estaciones del sur, que se encuentran en vecindarios más pobres y deteriorados de la ciudad.

Modelos de Machine Learning en el Análisis Inmobiliario

En los estudios realizados se han utilizado diferentes modelos de Machine Learning para medir la influencia de proyectos de infraestructura vial y transporte, enfocándose en los proyectos férreos, líneas de trenes y estaciones de trenes que impactan en la variación de los precios de bienes raíces en las áreas aledañas a su ubicación, aunque varios autores utilizan distancias entre 0 y 1 kilómetro para medir este impacto, se puede notar que entre más cerca se encuentre la propiedad de la estación de tren, la variación en el precio tiende a ser mayor.

Se han analizado diversos enfoques de Machine Learning que se utilizan para determinar el precio de las viviendas. Estos métodos se sustentan en referentes bibliográficos y datos históricos, destacándose las siguientes tendencias:

Modelos de Regresión

Regresión Hedónica. Modelo clásico que estima el valor de una vivienda a partir de sus características.

Regresión Lineal Múltiple. Técnica para modelar relaciones lineales entre el precio y múltiples variables explicativas.

Modelos Semiparamétricos de Regresión Espacial. Captura efectos espaciales no lineales.

Regresión Hedónica Espacial. Introduce la variable espacial para mejorar la precisión de los modelos de precios.

Modelos Avanzados

Gradient Boosting, LightGBM, y XGBoost. Algoritmos basados en boosting que mejoran la predicción mediante el ajuste de árboles de decisión secuenciales.

Árboles de Decisión. Técnica para la clasificación y regresión que segmenta los datos según valores de variables clave.

Aprendizaje Profundo de Regresión Supervisada. Redes neuronales aplicadas para capturar relaciones complejas entre variables.

Metodologías de Optimización y Evaluación

Algoritmos de Regresión en Grid Search. Optimización de hiperparámetros para mejorar el rendimiento de los modelos.

Revisión Literaria y Meta-análisis. Estudio comparativo y agregación de resultados de múltiples investigaciones.

Análisis Complementarios

Space Syntax. Análisis del espacio y su influencia en los patrones de comportamiento.

Prospectiva Territorial y Análisis de Tendencias. Herramientas para predecir el impacto de cambios territoriales en los precios.

Estimaciones de Panel Estáticas y Dinámicas. Modelos de datos longitudinales que analizan variaciones temporales.

Burbujas Inmobiliarias: Causas y Efectos

Una burbuja inmobiliaria es definida como un rápido aumento en el precio de los bienes inmuebles que generalmente viene acompañado por alcanzar niveles insostenibles, que luego de un tiempo y sin alguna razón aparente disminuyen de manera abrupta (Gómez, 2023). Entre las causas se encuentran alta demanda de las propiedades y poca oferta, lo que produce un alza en los precios de las viviendas; la especulación producida por el mismo mercado, donde las personas compran viviendas a precios bajos y las revenden a precios superiores, lo que produce el incremento de los precios ofertados, dicho fenómeno provoca también como consecuencia un aumento de los créditos en las entidades financieras, que si no se controla adecuadamente, se le otorgarían crédito a precios con baja capacidad de pago, de manera que no solo influye en el incremento de compra, sino que una vez se adquiere la vivienda, vienen los problemas en el sistema financiero para cumplir con la obligación adquirida.

Las burbujas inmobiliarias podrían llegar a afectar la economía de un país y por ende sobrevaluar o sobre costear los precios de las viviendas cuando realmente los precios son más bajos, influenciando a su vez el precio de los materiales de construcción, lo que conlleva a generar temor en el comprador y retrasar las decisiones de compra, afectando otros mercados por la presión en el sistema (Gómez, 2023).

Fenómenos de este tipo deben ser contemplados en la estimación de los precios de las viviendas bajo modelos de Machine Learning, ya que estos podrían estar influenciados no por las características físicas como tal de las viviendas y por los desarrollos adjuntos a las viviendas, sino también por comportamientos de este tipo y más cuando se trata de evaluar en este caso,

hasta qué punto podría llegar el precio de un inmueble considerando factores externos que no se confundan con burbujas inmobiliarias.

Contexto del Proyecto RegioTram de Occidente

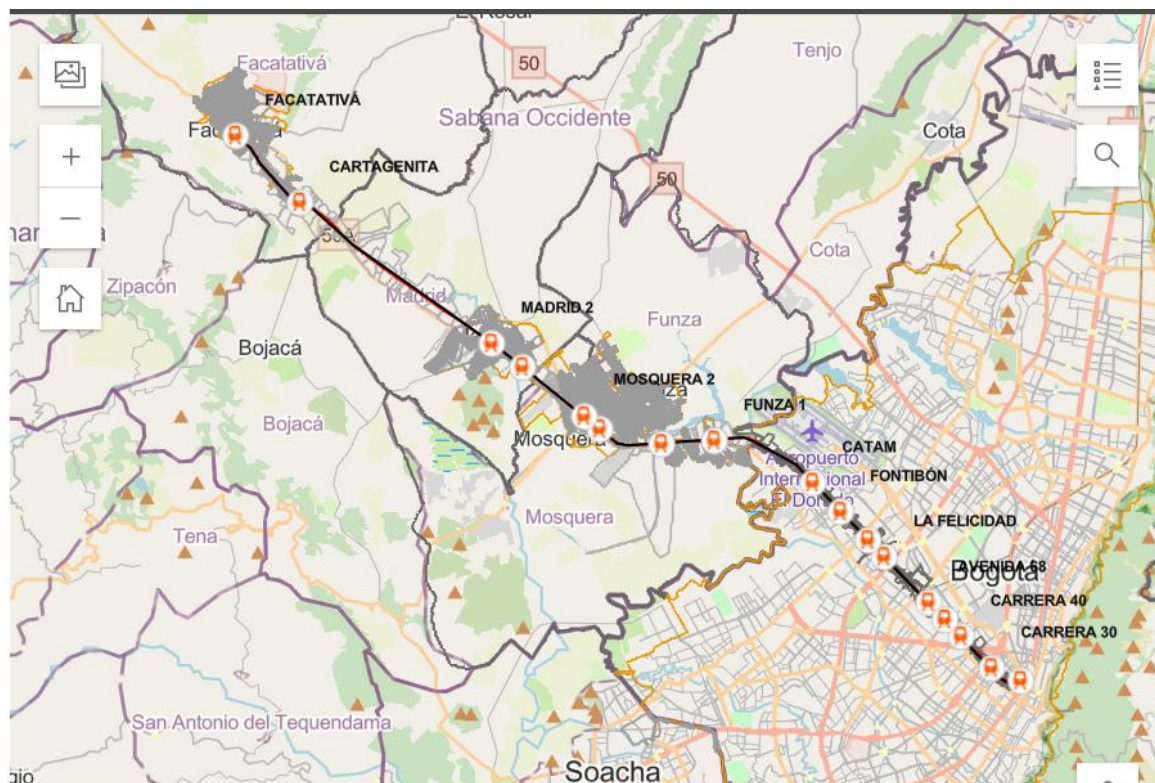
Descripción del Proyecto

El proyecto RegioTram de occidente es un tren eléctrico para el transporte de pasajeros regional, el cual conectará a los municipios de Funza, Mosquera, Madrid y Facatativá con el distrito capital (Bogotá), dicho proyecto se desarrollará en un corredor ferrero que ya existe, con una longitud aproximada de 40 km, de los cuales 25 km están distribuidos en los municipios de la Sabana y el resto (15 km) en Bogotá (Ministerio de Transporte, 2024).

El regioTram tal como se observa en la ilustración 1, está contemplado para arrancar en la carrera 17 con avenida calle 6 en la ciudad de Bogotá, en este punto se espera que conecte con la primera línea del Metro de Bogotá, posteriormente tomará la carrera 17 hacia el sur hasta la calle 20, y de allí la idea es que se conecte con el corredor Ferrero de occidente que ya existe, y desde ese punto cruza la NQS, la AV de las Américas, la Av. Carrera 50, la Av. Carrera 68, la Av. Boyacá, el Canal San Francisco y la Av. Ciudad de Cali, hasta la Estación Fontibón. Desde dicha estación continua por el corredor Férreo de Occidente hasta el límite del Distrito Capital con el municipio de Funza, luego Madrid, Mosquera y finalmente Facatativá – Calle 9, donde se localiza la estación final del trayecto.

Figura 1

Mapa de Trazado de RegioTram de Occidente



Nota. Proyectos férreos. Tomado de. Regiotram de occidente, por Ministerio de Transporte, (2024).

Dentro de las bondades que este proyecto contempla se encuentran la reducción de los tiempos de viaje a más de la mitad, conectando a Bogotá con Facatativá en menos de 55 minutos de trayecto, lo cual es una solución relevante para quienes viven o visitan los municipios o deben desplazarse diariamente a sus lugares de trabajo a Bogotá, puesto que actualmente los recorridos se hacen desde y hacia las calles 13 y 80, donde los viajeros pueden tardar más de 2 horas y 30 minutos para llegar a sus destinos. Adicional, este sistema tendrá la capacidad para movilizar a

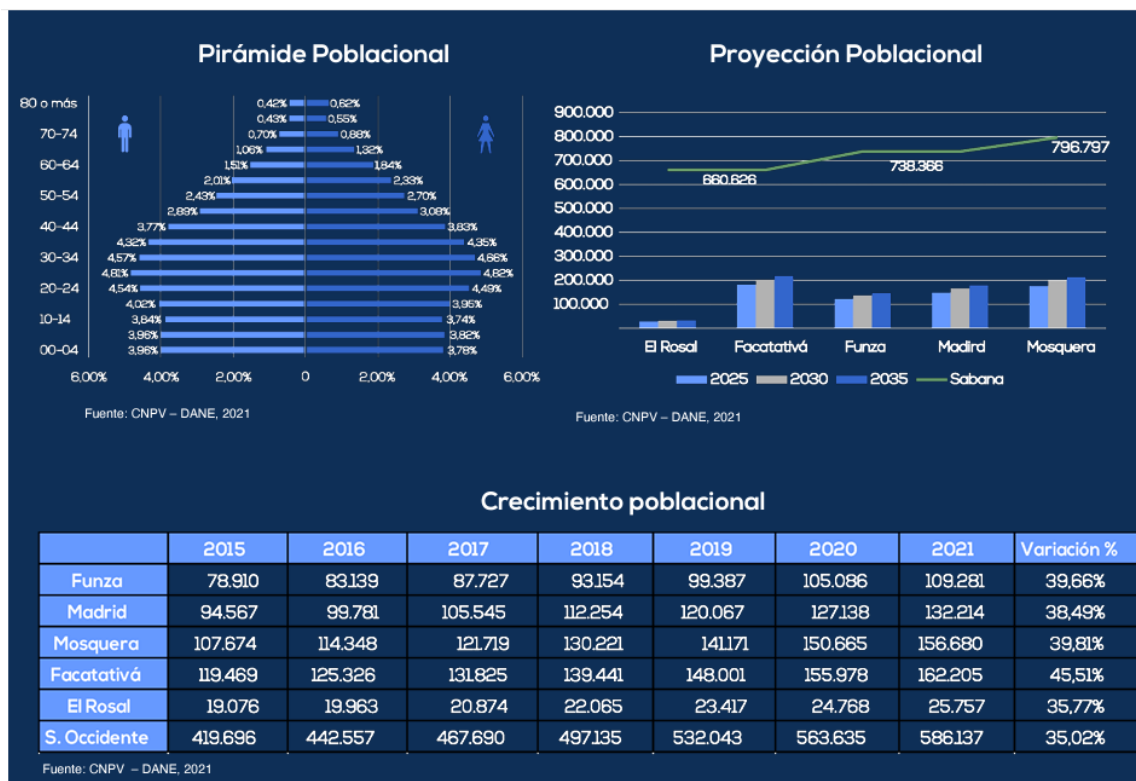
130.000 pasajeros por día, gracias a la flota de 18 trenes eléctricos que se moverán en los 40 kilómetros descritos anteriormente (Castillo, 2024).

Impacto Esperado en los Municipios Cercanos a Bogotá

Con respecto al proyecto del Regiotram de occidente que afectaría varios municipios cercanos a Bogotá, ciudad capital de Colombia y los estudios analizados relacionados con el impacto de diferentes desarrollos de infraestructura vial enfocándose en vías y estaciones de tren, se espera hacer un análisis de todas las variables que influyen en la variación de los precios de los inmuebles y de alguna manera determinar la influencia de desarrollos viales para predecir mediante modelos de Machine Learning la variación que podrían tener los precios de las propiedades residenciales y comerciales en áreas aledañas de no más de 1.5 kilómetros de distancia de la estación de tren por donde va a pasar el Regiotram de occidente.

Situación Actual del Mercado Inmobiliario en Mosquera, Madrid, Funza y Facatativá

De acuerdo con el estudio realizado por ProBogotá, (2022), la población tenderá a crecer en un 20.6% en los próximos 11 años, según las cifras del DANE, se pasará de 660.828 habitantes en 2025 a 796.797 habitantes en 2035, siendo Facatativá el municipio con mayor población y porcentaje de crecimiento. Adicional el 76.90% de la población se encuentra en edad de trabajar, lo que supone de acuerdo con la Figura 2, que no solo crecerá la población en términos de número de habitantes, sino que estos buscarán desplazarse dentro o fuera de los municipios para desarrollar sus diferentes actividades.

Figura 2*Población de la Sabana de Occidente*

Nota. Población y Proyección Poblacional. Tomado de. Plan Estratégico de Desarrollo para Sabana Occidente a 2051, por ProBogotá, 2022, Fundación para el progreso de la Región.

Adicional según las proyecciones del DANE, el número de personas por hogar va a tender a disminuir pasando de 2.99 a 2.39 (Tabla 2), resultado que impactará el número y tipología de vivienda requerida, sumado a que cerca del 34.70% de los hogares en estos municipios cuentan con ingresos menores o iguales a 1 SMMLV, según se puede evidenciar en la Tabla 3 y Tabla 4.

Tabla 2*Número de Personas por Hogar*

Número de personas por hogar	2020	2025	2030	2035
El Rosal	2.91	2.68	2.48	2.33
Facatativá	2.99	2.76	2.55	2.39
Funza	3.00	2.76	2.54	2.38
Madrid	3.03	2.79	2.58	2.41
Mosquera	3.04	2.80	2.59	2.42
Sabana Occidente	2.99	2.76	2.55	2.39

Nota. Promedio de personas por hogar. Tomado de. Plan Estratégico de Desarrollo para Sabana Occidente a 2051, por ProBogotá, 2022, Fundación para el progreso de la Región.

Tabla 3*Nivel de Ingresos por Hogar \$0 – \$3,5M*

Municipio	\$0 - \$0,8M	\$0,81M - \$1,50M	\$1,5M - \$2M	\$2M - \$2,5M	\$2,5M - \$3,5M
El Rosal	2.954	1.371	861	443	286
Facatativá	16.952	10.272	4.312	1.692	1.392
Funza	9.253	7.743	4.663	2.997	1.197
Madrid	10.454	10.458	5.351	2.135	1.481
Mosquera	12.518	10.481	5.649	2.569	3.648
S. Occidente	52.131	40.325	20.836	9.836	8.004

Nota. Ingresos por hogar. Tomado de. Plan Estratégico de Desarrollo para Sabana Occidente a 2051, por ProBogotá, 2022, Fundación para el progreso de la Región.

Tabla 4

Nivel de Ingresos por Hogar \$3,5M- Más de \$9M

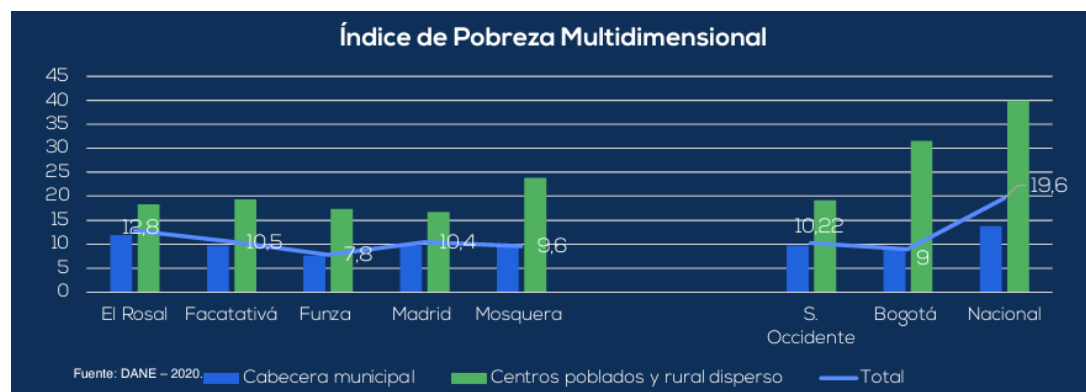
Municipio	\$3,5M - \$4,9M	\$5M - \$6,8M	\$6,7M - \$9M	Más De \$9M	NS/NR
El Rosal	31	24	31	-	404
Facatativá	493	-	155	201	5.398
Funza	977	500	553	384	949
Madrid	840	1.122	-	194	755
Mosquera	1.159	1.455	586	461	2.425
S. Occidente	3.500	3.101	1.325	1.240	9.931

Nota. Ingresos por hogar. Tomado de. Plan Estratégico de Desarrollo para Sabana Occidente a 2051, por ProBogotá, 2022, Fundación para el progreso de la Región.

Por su parte, los índices de pobreza multidimensional en los municipios de la sabana de occidente se encuentran entre 7% y 13%, comparado con Bogotá que se ubica en un promedio del 9% (Figura 3), donde Mosquera es de los municipios que ha recibido mayor saldo migratorio, según se observa en la Figura 4.

Figura 3

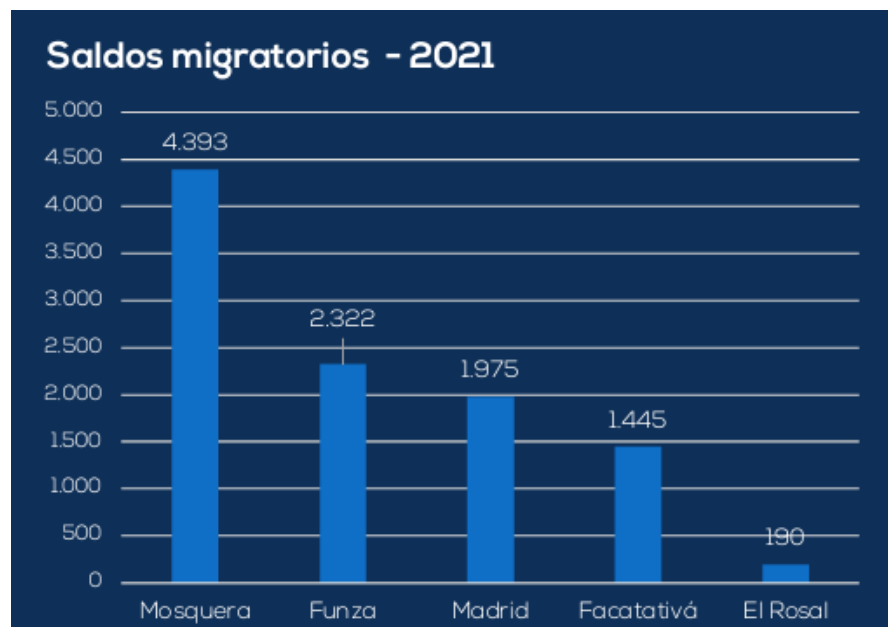
Condiciones de Pobreza



Nota. Índice de Pobreza Multidimensional. Tomado de. Plan Estratégico de Desarrollo para Sabana Occidente a 2051, por ProBogotá, 2022, Fundación para el progreso de la Región.

Figura 4

Saldos Migratorios -2021



Nota. Saldos migratorios. Tomado de. Plan Estratégico de Desarrollo para Sabana Occidente a 2051, por ProBogotá, 2022, Fundación para el progreso de la Región.

Finalmente, las dinámicas poblacionales, sociales y económicas que se describen en las tablas e ilustraciones anteriores, dan cuenta de un acelerado crecimiento en la producción de vivienda en los últimos 10 años, con un incremento del 144.4% de las viviendas, tal como se muestra en la Figura 5.

Figura 5

Vivienda



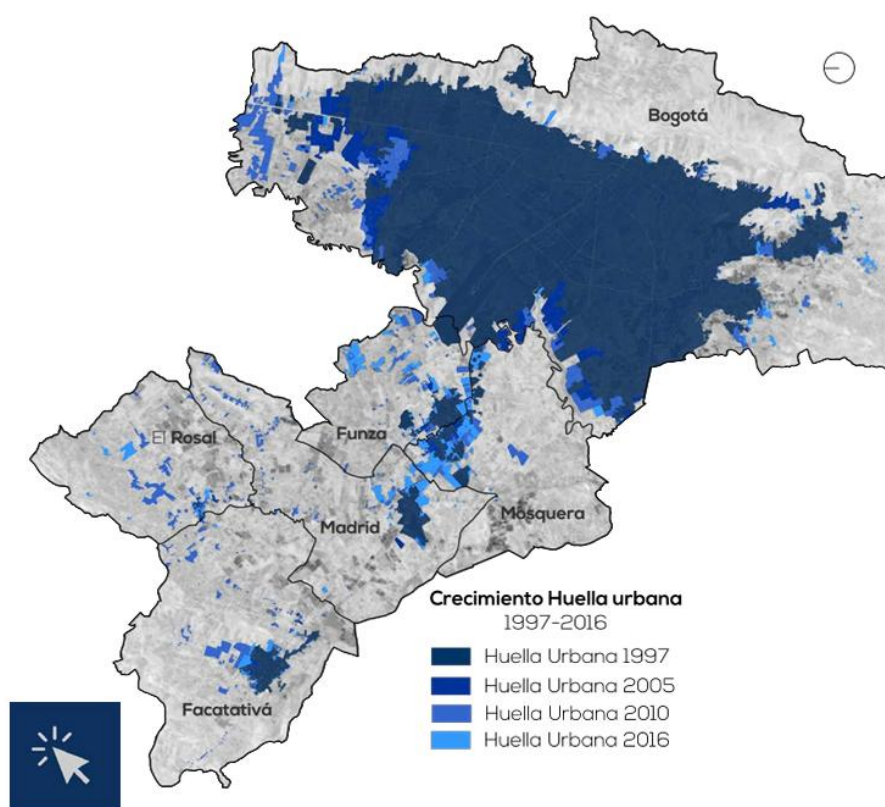
Nota. Crecimiento unidades de vivienda. Tomado de. Plan Estratégico de Desarrollo para Sabana Occidente a 2051, por ProBogotá, 2022, Fundación para el progreso de la Región.

Producto de la ubicación estratégica también es importante mencionar que, por su localización, por la cercanía con los dos centros logísticos y de distribución más grandes del país como lo son el Aeropuerto El Dorado y Corabastos, una distancia aproximada de 30 kilómetros, dicha industria concentra la generación de empleo, que contribuye en un 20% del valor agregado del Departamento de Cundinamarca. Sin embargo, se evidencia insuficiencia en las conexiones viales con Bogotá sobre la Calle 80 y la Calle 13, lo cual se prevé que pueda ser subsanado con la entrada en operación del Regiotram para el año 2028, que permitirá la conexión entre Facatativá y Bogotá, pues hoy se habla de cerca de 550.000 viajes que son realizados de los diferentes municipios de la Sabana hacia Bogotá, donde hay una creciente dinámica de viajes especialmente desde Mosquera y Funza, donde el promedio del viaje es de 84 minutos.

Ahora bien, a raíz del acelerado crecimiento urbano de la Sabana, se generó una fuerte presión sobre el suelo de los municipios, lo que implicó nuevos retos en la planificación para las entidades territoriales, producto de un crecimiento casi que poco controlado de la capital Bogotá, como se evidencia en la Figura 6, la fuerte presión en el uso del suelo de la ciudad, ha obligado de cierta manera a que las personas se desplacen hacia los municipios, sumado a que ha habido un crecimiento industrial a los largo de la calle 13, evidenciando una mezcla en los usos del suelo, con deficiencia en espacio público y por ende en la reducción de los espacios para construir más viviendas en Bogotá.

Figura 6

Crecimiento de la Huella Urbana



Nota. Huella urbana. Tomado de. Plan Estratégico de Desarrollo para Sabana Occidente a 2051, por ProBogotá, 2022, Fundación para el progreso de la Región.

Lo anterior, ha generado presión en la Sabana. La huella urbana generada por la ciudad requiere alrededor de 15.000 hectáreas, donde no hay espacio y por lo cual requerirá la búsqueda de 3.113 hectáreas para construir fuera de la ciudad. Por lo tanto, según las proyecciones realizados por el DANE se espera que se tenga un 66.04% de aumento en el número de unidades habitacionales ofertadas y que dicho incremento sea cubierto por la sabana de occidente.

Modelos de Machine Learning para Analizar el Impacto del Regiotram de Occidente en la Valorización Inmobiliaria en sus Áreas Cercanas

Con el fin de realizar el análisis de los modelos de machine learning, nos enfocamos en indagar cómo el desarrollo del Regiotram de Occidente afecta la valorización de las propiedades cercanas a sus estaciones mediante la implementación de modelos de machine learning. Partimos de una selección rigurosa de variables relevantes que permiten capturar tanto las características intrínsecas de las propiedades como factores externos clave, como la proximidad al sistema de transporte y las condiciones del entorno. Este análisis tiene como objetivo principal identificar patrones que expliquen cómo los proyectos de infraestructura vial influyen en el mercado inmobiliario, brindando una herramienta robusta para la toma de decisiones basada en datos.

Para garantizar la calidad y pertinencia de este estudio, seguimos criterios estrictos en la selección de referencias bibliográficas y metodológicas. Elegimos investigaciones recientes y académicamente reconocidas que aplican enfoques estadísticos y técnicas de machine learning relacionadas con el análisis de precios de vivienda e infraestructura vial. Esta base nos permitió definir como variable objetivo el precio de la propiedad, representado como un valor continuo ya que es el punto central del estudio y mide directamente el efecto del Regiotram en el mercado inmobiliario. Mientras que las variables independientes incluyeron aspectos como área, antigüedad, número de habitaciones y baños, y la distancia a las estaciones del Regiotram. Este conjunto de variables aporta una visión integral que facilita una comprensión más profunda del fenómeno estudiado.

Adicionalmente, exploramos tres modelos de machine learning, regresión lineal múltiple, árboles de decisión y redes neuronales supervisadas. Cada modelo se seleccionó por sus características particulares; por ejemplo, la regresión lineal permite interpretaciones claras, es

una herramienta estadística que nos ayuda a identificar relaciones lineales entre las variables independientes y el precio de las propiedades. Este modelo tiene como ventaja su facilidad de interpretación, ya que ofrece coeficientes claros que explican el impacto de cada variable. Sin embargo, presenta limitaciones cuando las relaciones entre variables son no lineales o complejas, lo que reduce su capacidad para capturar todas las dinámicas presentes en los datos, mientras que los árboles de decisión son útiles para detectar relaciones no lineales, estos ofrecen la capacidad de identificar relaciones no lineales entre las variables, nos brinda un enfoque es ideal para problemas donde las relaciones entre los datos no son evidentes, además de ser fácil de visualizar y robusto frente a valores atípicos. No obstante, los árboles de decisión presentan el riesgo de sobreajuste, especialmente cuando el conjunto de datos es limitado, y pueden ser menos precisos en problemas muy complejos.

Finalmente, revisamos el modelo de redes neuronales supervisadas, el cual ofrece un enfoque de aprendizaje profundo que es especialmente útil para manejar conjuntos de datos grandes y detectar patrones no lineales complejos. Este modelo es capaz de identificar relaciones ocultas en los datos que otros métodos podrían pasar por alto. Sin embargo, tiene desventajas significativas, como la necesidad de mayores recursos computacionales y tiempo de entrenamiento. Además, requiere un volumen considerable de datos para evitar el sobreajuste, lo que puede ser un desafío en estudios con conjuntos de datos más pequeños. Las redes neuronales destacan en su capacidad para manejar datos complejos. A lo largo del capítulo, detallamos las ventajas y desventajas de cada enfoque, proporcionando un análisis crítico que permite evaluar cuál se ajusta mejor a los datos disponibles y al objetivo del estudio.

Diseño del Estudio

Se ha definido los siguientes criterios de selección para llevar a cabo el estudio:

Relevancia: Las tesis deben estar directamente relacionadas con el tema de investigación, en este caso, análisis de precios de viviendas, proyectos de infraestructura vial y machine learning.

Calidad Académica: Priorizar tesis de universidades o instituciones reconocidas que sigan estándares académicos rigurosos.

Actualidad: Asegurar que las fuentes sean recientes, preferentemente publicadas en los últimos 5 a 10 años, para reflejar tendencias y tecnologías actuales, como el uso de machine learning.

Metodología: Selecciona tesis que utilicen metodologías robustas, como análisis estadístico, econometría, o machine learning, ya que son afines al enfoque de la monografía.

Para este análisis, definimos como variable objetivo el precio de la propiedad, representado como un valor continuo. Esta variable es el punto central del estudio, ya que mide directamente el efecto del Regiotram en el mercado inmobiliario. Por otro lado, seleccionamos las variables independientes en función de que permiten capturar tanto las características intrínsecas de las propiedades como los elementos externos que influyen en su valoración. Entre estas variables se incluyó la proximidad al Regiotram, que mide la distancia de cada propiedad a la estación más cercana. También consideramos características internas de las propiedades, como el área construida, la antigüedad, el número de habitaciones, baños y estacionamientos, y características del entorno, como la cercanía a escuelas, hospitales, centros comerciales y la accesibilidad vial. Estas variables aportan una visión integral de los factores que influyen en el valor de las viviendas.

Para modelar estas relaciones, analizamos tres técnicas principales de machine learning. En primer lugar, utilizamos la regresión lineal múltiple, una herramienta estadística que permite

identificar relaciones lineales entre las variables independientes y el precio de las propiedades. Este modelo tiene como ventaja su facilidad de interpretación, ya que ofrece coeficientes claros que explican el impacto de cada variable. Sin embargo, presenta limitaciones cuando las relaciones entre variables son no lineales o complejas, lo que reduce su capacidad para capturar todas las dinámicas presentes en los datos.

Como segundo modelo, encontramos árboles de decisión, que ofrecen la capacidad de identificar relaciones no lineales entre las variables. Este enfoque es ideal para problemas donde las relaciones entre los datos no son evidentes, además de ser fácil de visualizar y robusto frente a valores atípicos. No obstante, los árboles de decisión presentan el riesgo de sobreajuste, especialmente cuando el conjunto de datos es limitado, y pueden ser menos precisos en problemas muy complejos.

Finalmente, revisamos el modelo de redes neuronales supervisadas, un enfoque de aprendizaje profundo que es especialmente útil para manejar conjuntos de datos grandes y detectar patrones no lineales complejos. Este modelo es capaz de identificar relaciones ocultas en los datos que otros métodos podrían pasar por alto. Sin embargo, tiene desventajas significativas, como la necesidad de mayores recursos computacionales y tiempo de entrenamiento. Además, requiere un volumen considerable de datos para evitar el sobreajuste, lo que puede ser un desafío en estudios con conjuntos de datos más pequeños.

Recolección de Datos

La recolección de datos para el análisis predictivo del impacto del RegioTram de Occidente en los precios de vivienda fue realizada mediante un enfoque meticuloso y estructurado, basado en fuentes primarias y secundarias de alta relevancia. Se empleó una metodología monográfica que incluyó la revisión exhaustiva de estudios, artículos científicos y

bases de datos catastrales. Se investigaron tanto las dinámicas del mercado inmobiliario en los municipios de Mosquera, Madrid, Funza y Facatativá en Cundinamarca, como estudios previos sobre la relación entre infraestructuras viales y valorización inmobiliaria en distintas partes del mundo, lo cual permitió establecer comparaciones contextuales y enriquecer el marco teórico de la investigación. Ejemplos de fuentes incluyen el Observatorio Técnico Catastral y el Departamento Administrativo Nacional de Estadística (DANE), que proporcionaron datos cuantitativos actualizados sobre precios de vivienda, características demográficas y estratos socioeconómicos de la región.

La selección de fuentes para esta investigación siguió criterios claros y estrictos para garantizar la calidad y relevancia de los datos utilizados. En primer lugar, se priorizó la relevancia, seleccionando únicamente estudios que abordaran temas relacionados directamente con los precios de vivienda, proyectos de infraestructura vial y la aplicación de modelos de machine learning. Asimismo, se valoró la calidad académica, escogiendo tesis e investigaciones provenientes de universidades e instituciones reconocidas que cumplieran con estándares académicos rigurosos. Para asegurar que los hallazgos reflejaran las tecnologías y tendencias actuales, también se consideró la actualidad de las fuentes, dando preferencia a publicaciones realizadas en los últimos cinco a diez años. Finalmente, se seleccionaron estudios que aplicaran metodologías robustas, como análisis estadísticos, econometría y machine learning, ya que estas herramientas son esenciales para el enfoque técnico y predictivo de este trabajo.

La inclusión de estos criterios permitió desarrollar un marco metodológico que no solo se apoyó en investigaciones previas relevantes, sino que también identificó patrones replicables en casos internacionales. Por ejemplo, estudios realizados en Phoenix y Teherán revelaron que la proximidad a sistemas ferroviarios incrementa el valor de las propiedades residenciales y

comerciales, resultados que fueron corroborados por investigaciones de Debrezion et al., (2007) y Tyagi y Singhal, (2018). Estos trabajos demostraron cómo la accesibilidad al transporte público puede influir significativamente en la valorización de los inmuebles, un fenómeno que, adaptado al caso del RegioTram, permite comprender las dinámicas de precios esperadas en la región. Además, el análisis de tesis recientes destacó la importancia de considerar variables adicionales, como el contexto socioeconómico y la infraestructura urbana, para obtener un entendimiento más completo del impacto del proyecto.

La metodología combinó fuentes cuantitativas y cualitativas, lo que permitió un análisis integral. En el ámbito cuantitativo, se recopilaron datos sobre precios del metro cuadrado y características de las propiedades en los municipios aledaños al proyecto, utilizando bases de datos oficiales y reportes inmobiliarios. Por otro lado, el componente cualitativo incluyó estudios de percepción ciudadana y expectativas sobre la valorización en las zonas cercanas al RegioTram, lo que permitió contextualizar los resultados y entender las implicaciones del proyecto desde una perspectiva social. Estas aproximaciones permitieron identificar tendencias clave y relacionarlas con factores como la proximidad a las estaciones proyectadas, la accesibilidad al transporte y la disponibilidad de servicios educativos y comerciales en la región.

Finalmente, las variables seleccionadas para el análisis predictivo incluyeron elementos fundamentales como la distancia a las estaciones del RegioTram, precios históricos de las propiedades, niveles de accesibilidad y características socioeconómicas. Estas variables se integraron en modelos de machine learning avanzados, como regresiones espaciales y análisis hedónicos, que son herramientas reconocidas en la literatura para prever tendencias en los precios de propiedades. Este enfoque, sustentado en estudios recientes como los de García, (2021) y Abidoje et al., (2021), garantizó una base sólida para analizar el impacto del

RegioTram de Occidente en el mercado inmobiliario de la Sabana de Occidente, ofreciendo predicciones fundamentadas que reflejan las dinámicas del proyecto en el contexto local e internacional.

Variables Analizadas

Con el fin de analizar la influencia de la implementación del Regiotram de occidente en los municipios de Mosquera, Funza, Madrid y Facatativá, los cuales son objeto de este estudio y luego de recopilar las técnicas de machine learning más populares, se identificó que la literatura clasifica la información tanto de tipo cualitativa como cuantitativa, como determinantes para predecir el precio de las viviendas.

Según Munevar, (2016), los inmuebles se consideran mercancías con características muy particulares, lo que hace que se diferencien de otros bienes por tal razón la valoración de los inmuebles se da en función directa de las siguientes variables:

Factor área: delimitación geográfica de la zona de estudio.

Factor edad: antigüedad del inmueble.

Factor piso de ubicación: piso de ubicación dentro del edificio, de tratarse de un conjunto residencial.

Factor número de garajes: número de garajes con los que cuenta la unidad.

Por su parte Abdul-Rahman et al., (2024), aseguran que el precio de las viviendas se ve afectado significativamente por diversos factores para determinar un precio razonable de la vivienda, lo que implica un proceso de cálculo del mismo. Sin embargo, para efectos de la determinación de los precios, las variables más usadas son:

Precio: como variable objetivo, de tipo continuo y que corresponde al precio del inmueble vendido en el mercado.

Ubicación: es una variable de tipo nominal y corresponde a la ubicación geográfica del inmueble, es decir, la zona o región.

Habitación: es una variable de tipo discreto y corresponde al número de habitaciones que tiene el inmueble.

Baño: es una variable de tipo discreto y corresponde al número de baños que tiene el inmueble.

Aparcamiento: es una variable de tipo discreto y corresponde al número de plazas de aparcamiento o estacionamientos (garajes) del inmueble.

Tamaño: es una variable de tipo continuo y corresponde a la superficie construida de una casa o la superficie del terreno de una casa.

Mobiliario: es una variable de tipo nominal y corresponde al estado del mobiliario de una casa, tales como completamente amueblado, semi amueblado o sin amueblar.

Tipo de propiedad: es una variable de tipo nominal y corresponde al tipo de condominio, tales como esquina, dúplex, lote final, intermedio, ático, estudio o triple.

Centro comercial: es una variable de tipo continuo y corresponde al centro comercial más cercano al inmueble, medido en número de kilómetros.

Escuela: es una variable de tipo continuo y corresponde a la escuela más cercana al inmueble, medido en número de kilómetros.

Hospital: es una variable de tipo continuo y corresponde al hospital más cercano al inmueble, medido en número de kilómetros de distancia.

Tren de metro: es una variable de tipo continuo y corresponde a la estación de tren más cercano al inmueble, medido en número de kilómetros de distancia.

Sin embargo, hay quienes afirman que el sector inmobiliario está evolucionando y cambiando rápidamente, porque en las últimas décadas se ha visto un aumento de la demanda de viviendas y cada día tienden a aparecer proyectos de viviendas de lujo a precios exequibles (Özögür et al., 2023). Es por esta razón que la fijación de precios dinámicos que estime los precios de las viviendas utilizando las características más importantes de las casas se hace fundamental que simplemente el área de metros cuadrados que pueda tener. Por lo anterior, las variables que se deberían tener en cuenta según estos autores son:

Metro cuadrado: como variable de tipo numérico indica cuantos metros cuadrados tiene el inmueble.

Habitación: variable de tipo número que indica cuantas habitaciones tiene el inmueble.

Baños: variable de tipo numérico que indica el número total de baños en el inmueble.

Piso de planta: variable de tipo numérico que indica el número de plantas o pisos con las que cuenta el inmueble.

Pisos en el edificio: variable de tipo numérico que indica el número total de pisos del edificio.

Dentro del sitio: variable de tipo categórica que clasifica la vivienda en 0 y 1, dependiendo si se encuentra en un conjunto residencial o no.

Fuera del sitio: variable de tipo categórica que clasifica la vivienda en 0 y 1, dependiendo si a menos de 10 kilómetros de distancia se encuentra un centro de salud o una escuela o una estación para tomar transporte público.

Año: variable de tipo categórica que clasifica la vivienda en 0 y 1 para dos años diferentes de venta.

Precio: variable de tipo numérico y objetivo, que permite determinar el precio de venta del inmueble.

Finalmente, Tyagi y Singhal, (2018), identificaron basados en un caso de estudio de las viviendas de Delhi, que el ferrocarril mejora significativamente el valor de la tierra y la propiedad alrededor de las estaciones debido a la mejor accesibilidad, de hecho el metro desarrollado en Hong Kong da muestra de este indicio, donde la captura del valor de la tierra y la propiedad logran un desarrollo urbano de la ciudad, capturando el valor de la tierra, estableciéndose como una solución exitosa a nivel mundial y de hecho identificándolo como potencial para la financiación de la infraestructura ferroviaria urbana, en este contexto, los valores se capturan mediante la monetización del aumento en los precios de los inmuebles debido al efecto directo de la infraestructura.

En este sentido, vivir más cerca de este tipo de instalaciones ha reducido el costo de los traslados o desplazamientos y ha mejorado la accesibilidad a expensas de mayores precios de la tierra e incluso de alquileres, es por esta razón que las variables utilizadas son:

La proximidad a las estaciones del metro en términos de kilómetros de distancia.

Los precios de las propiedades comerciales.

El tiempo dividido en dos periodos: 2000-2004 (planificación y construcción) y 2005 a 2008 operación.

Las características intrínsecas de la vivienda tales como número de metros cuadrados, número de habitaciones, número de baños y número de estacionamientos.

Aplicación de Modelos de Machine Learning

Ventajas y Desventajas de los Modelos

Regresión Lineal Múltiple. Útil en estudios exploratorios iniciales, pero insuficiente para complejidad de datos. Se consideran como ventajas el uso de regresión lineal múltiple para predecir el precio de una vivienda cuando es de fácil interpretación, es decir, los coeficientes de la regresión muestran como cada variable independiente (área de la vivienda, número de habitaciones, baños, estacionamientos, etc.) afecta el precio de la vivienda, debido a que parte de una función del tipo: $y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \dots + \beta_k * X_k$, donde y es la variable dependiente, que para este caso es el precio de la vivienda, β_0 es la constante o intersección con el eje de las X_1, X_2, X_k , es decir, las variables explicativas y β_1, β_2, \dots y β_k , que corresponden al coeficiente asociado a cada una de las variables, el cual da la razón de los cambios de las distintas variables explicativas cuando se presenta una variación en una unidad, lo que la hace muy eficiente en términos de procesamiento, adecuada para conjuntos de datos pequeños o medianos, porque incluso se puede desarrollar en programas como Excel.

Adicional es importante mencionar que este tipo de modelos ayuda a identificar relaciones entre las variables, determinando cuáles factores tienen un impacto significativo en el precio de la vivienda, ya que permite establecer la colinealidad de las variables, es decir, la identificación de una situación donde exista alta correlación entre las variables existentes, que se mide a través del coeficiente de correlación R^2 , que va de cero a uno o p-valores para evaluar la calidad y significancia del modelo, de manera tal que, se adapta fácilmente para incluir más variables predictoras (Munevar, 2016).

Sin embargo, este tipo de modelos también presenta sus desventajas, entre las cuales se mencionan la sensibilidad a la multicolinealidad, la cual se presenta cuando las variables

independientes están altamente correlacionadas entre sí, afectando la precisión de los coeficientes, además de que este tipo de modelos asume relaciones lineales y por tanto puede tener dificultades para capturar relaciones dinámicas, tales como datos estacionales o fluctuaciones del mercado, siendo muy sensible a valores atípicos, datos faltantes o mal etiquetados y presentando riesgo de sobreajuste o subajuste en los modelos (Munevar, 2016).

Algoritmo Híbrido. Buen balance entre interpretación y precisión, aunque puede requerir ajustes para evitar sobreajuste. El algoritmo híbrido propuesto por Özögür et al., (2023), que combina técnicas de regresión hedónica y aprendizaje automático, incluidos métodos de agrupamiento y clasificación utiliza la salida de un algoritmo como entrada de otro dentro de todo el esquema, de manera que divide los datos de entrenamiento en grupos incorporando los errores de la regresión múltiple en el algoritmo k-means, realizando predicciones de clase para datos de prueba usando k-nn, donde se utiliza las etiquetas de clase como variable dependientes, estimando el precio de cada clúster identificado usando el SVR, la cual bajo esta técnica los autores son los pioneros en la fijación de precios de viviendas con el uso de algoritmos híbridos.

Como vimos en la sección anterior el modelo de regresión lineal tiene bondades, pero también múltiples desventajas, es por esta razón que combinando varios métodos se logra tener un buen resultado. Por su parte la agrupación de K-medias es un método que agrupa objetos similares minimizando la distancia dentro del clúster y maximizando la distancia entre los clúster, razón por la cual es uno de los métodos más utilizados, ya que maximiza la distancia entre los centroides y las observaciones dentro del clúster. El método K-NN si bien es un algoritmo simple que se utiliza tanto en la regresión como en la clasificación, sus bondades giran en torno a que su tiempo de cálculo es bajo y su interpretación es sencilla de analizar, para efectos del uso que se le dio a este algoritmo en el estudio de la predicción de los precios de las

viviendas, es importante mencionar que predice las etiquetas de clase de los datos de prueba al observar los datos de entrenamiento y finalmente bajo el modelo de regresión de vectores de soporte (SVR) a diferencia de una regresión lineal, este método no lo es, en el sentido de interpretar los puntos de datos en el espacio de salida utilizando funciones de kernel, lo que lo hace popular porque no tiene ningún supuesto para los datos.

Ahora bien, el algoritmo híbrido desarrollado donde los resultados de cada tomador de decisiones se utiliza como entrada para el siguiente paso, permite realizar correcciones en las observaciones mal clasificadas, teniendo como una de las principales ventajas precisamente la corrección dentro de los puntos de datos de entrenamiento, el algoritmo que se propone se puede utilizar para configurar un sistema dinámico que determinará automáticamente el valor de las residencias utilizando sus propiedades; dando como resultado que el enfoque si se puede utilizar para crear un sistema dinámico para la predicción de precios de viviendas basado en el aprendizaje estadístico, debido a que tiene la capacidad de actualizar las predicciones con el paso de corrección que aplica para una nueva observación adicional al conjunto de entrenamiento y por tanto, es configurable para realizar una estimación automática del precio de la vivienda procesando cada nueva información ingresada. En conclusión, al utilizar este método híbrido se minimizan los errores innatos de cada modelo, porque la salida de uno es la entrada secuencial del otro, lo que lo hace estadísticamente más robusto.

XGBoost. Ideal para predicciones precisas en problemas con múltiples variables, aunque menos eficiente si los datos son limitados. El XGBoost es considerado el algoritmo más poderoso para cualquier problema de regresión y clasificación, ya que incorpora árboles de decisión escalables y maneja tanto variables numéricas como categóricas, según lo menciona Abdul-Rahman et al., (2024), la mayoría de los estudios donde se ha hecho uso de este tipo de

algoritmo, este modelo es capaz de proporcionar el valor RMSE más bajo para la predicción de los precios de las viviendas en comparación con otros modelos.

Adicional, además de ser altamente efectivo en tareas predictivas, tiene la enorme ventaja y es que tiene la capacidad de manejar relaciones no lineales, incorpora técnicas para tratar valores faltantes, lo que reduce la necesidad de preprocesamiento extenso de los datos, computacionalmente es óptimo y aunque no es tan interpretable como un modelo de regresión lineal, permite identificar la importancia de las variables predictoras. Sin embargo, entre sus desventajas también se puede mencionar que aplicar un modelo XGBoost requiere un mayor conocimiento técnico y ajustes de hiperparámetros para obtener un buen rendimiento con el fin de evitar los sobreajustes, así como procesar grandes volúmenes de datos requerirá más recursos computacionales ya que al evaluarse en conjuntos de datos pequeños no es tan efectivo.

Comparación con Estudios Previos

Kuala Lumpur. Modelos avanzados como XGBoost mostraron mejor rendimiento al capturar relaciones complejas, ya que muestra el rango de las características en secuencia de mayor a menor: tamaño de la casa, número de garajes o estacionamientos y la distancia a alguna característica de transporte público, dicho modelo arrojó un R-cuadrado más alto en comparación con otros modelos, lo que lo hizo capaz de predecir los precios de las viviendas con el mejor valor de coeficiente de determinación, logrando generar resultados que fueron más consistentes y justificables ya que proporcionó una mejor capacidad predictiva, con el menor error medio absoluto y el menor error cuadrático medio.

Delhi. La proximidad al metro demostró un impacto significativo, especialmente tras la operación, con mejores resultados utilizando modelos lineales, como la regresión múltiple, demostrando que existe un efecto positivo del metro de Delhi en los precios de venta de las

unidades de vivienda residencial en estaciones seleccionadas del corredor del metro. De hecho, la proximidad al sistema ferroviario del metro mostró un efecto positivo en los precios de las propiedades, lo que la hace generar oportunidades significativas para la evaluación de precios en áreas de desarrollo. Los resultados de este estudio revelan que la estación del metro tiene un efecto de beneficio para las partes del sur de Delhi, donde en efecto los precios de las viviendas son más elevados.

Conclusión y Modelo Recomendado

Con base en los resultados esperados y los estudios previos, se recomienda usar modelos de regresión lineal múltiple si de datos pequeños se trata y contemplando relaciones lineales, con las precauciones ya descritas anteriormente. Sin embargo, un estudio juicioso aterrizado más a la realidad de un caso como la influencia del Regiotram en los precios de las viviendas de los municipios de Funza, Mosquera, Madrid y Facatativá requiere de métodos como Gradient Boosting (ej., XGBoost) para manejar relaciones no lineales y complejidad de datos, ya que este proporciona equilibrio entre precisión y facilidad de interpretación, adaptándose bien al impacto del Regiotram en diversas fases, de manera que es ideal cuando se busca un alto rendimiento predictivo y se dispone de datos grandes y variados, que no necesariamente son lineales, y que permiten evaluar el contexto, los datos y los recursos disponibles.

Resultados

Predicciones de Variación en los Precios de las Viviendas

El análisis del impacto del Regiotram de Occidente prevé un incremento en los precios de las propiedades cercanas a sus estaciones. La construcción de este sistema ferroviario representa un atractivo tanto para inversionistas como para compradores potenciales, ya que mejora sustancialmente la conexión entre Bogotá y los municipios de la Sabana de Occidente. En otras ciudades del mundo, se ha documentado cómo la proximidad a nuevas infraestructuras de transporte impulsa el valor de los inmuebles, y se espera un comportamiento similar en los municipios de Funza, Mosquera, Madrid y Facatativá. Según el trabajo desarrollado, se estima que las propiedades ubicadas a menos de 1.5 kilómetros de las estaciones experimenten un aumento en su valor, lo cual está respaldado por estudios en contextos comparables que registran aumentos de entre el 10% y el 20% en los precios de viviendas cercanas a infraestructuras ferroviarias (Castillo, 2024).

Esta predicción está fundamentada en estudios similares con técnicas de machine learning aplicadas a precios de vivienda en ciudades de todo el mundo. En particular, el uso de modelos de regresión y técnicas avanzadas de inteligencia artificial han mostrado que variables como la distancia a estaciones de tren, la frecuencia de servicios y el tiempo de desplazamiento a centros urbanos son determinantes en el precio de propiedades. Así, en el caso del Regiotram, se espera que su implementación cree un “efecto valorizador” que aumente el atractivo y el valor del suelo urbano en estas áreas de Cundinamarca, sobre todo, si tenemos en cuenta que al conectar la capital del país con estos municipios, esto impulse el movimiento de capital de las personas que quieran salir de la ciudad y pasar un tiempo o el fin de semana en un pueblo sin

tanta contaminación y congestión, al aumentar la demanda de propiedades, el precio de estas puede incrementar en poco tiempo.

Identificación de las Variables Clave: Transporte y Precios Inmobiliarios

Las variables clave que afectan la relación entre el transporte y los precios inmobiliarios son: la proximidad a la estación, la frecuencia y accesibilidad del servicio, y el entorno socioeconómico. Estos factores se combinan con características propias de cada municipio, como su estrato socioeconómico, la infraestructura educativa y comercial, y los servicios de salud, los cuales influyen directamente en la percepción de valor del área.

El análisis de variables en estudios previos muestra que las propiedades cercanas a estaciones de tren suelen registrar un incremento en su valor, pero este efecto varía dependiendo de aspectos como el tipo de zona (urbana, suburbana, o rural) y la calidad del transporte. Por ejemplo, Debrezion et al., (2007) destacan que el valor de las propiedades aumenta significativamente en las proximidades de estaciones de ferrocarril que facilitan el acceso a zonas comerciales y de trabajo en áreas metropolitanas. Esta investigación subraya que la proximidad a nodos de transporte no solo mejora el valor económico de las viviendas, sino que también puede elevar el perfil socioeconómico de las comunidades aledañas, haciendo del Regiotram una inversión estratégica en términos de desarrollo urbana.

Comparación de Resultados con Estudios Previos

Los resultados de este estudio del Regiotram de Occidente se comparan con investigaciones en otros proyectos de infraestructura de transporte masivo, como el Metro de Medellín y el Tren de Cercanías de Madrid. Estos casos presentan evidencia de cómo el transporte de alta calidad aumenta la demanda inmobiliaria en áreas antes menos valoradas. Por ejemplo, el impacto del Metro de Medellín en zonas suburbanas demostró que el valor de las

propiedades incrementó en un 12% a 15% en zonas conectadas directamente a las estaciones, especialmente donde antes había limitado acceso al centro urbano.

En el caso del Regiotram, se proyecta un efecto similar, argumentando que los municipios de la Sabana de Occidente tienen un perfil de alta movilidad hacia Bogotá y dependen de un transporte eficiente para facilitar el acceso laboral. Además, el documento cita casos en los que el incremento en precios se distribuye de manera heterogénea, mostrando un mayor impacto en zonas de estrato medio y bajo, donde el acceso a transporte puede transformar la accesibilidad al empleo y a servicios urbanos, una tendencia que podría repetirse en Cundinamarca.

Discusión

Impacto del Regiotram de Occidente en el Mercado Inmobiliario

El Regiotram de Occidente no solo mejorará la conectividad en Cundinamarca, sino que también transformará el mercado inmobiliario en los municipios cercanos, impulsando la valorización de propiedades residenciales y comerciales. La cercanía a Bogotá y la reducción de tiempos de desplazamiento se traducen en un atractivo adicional para los compradores e inversionistas. Este proyecto tiene el potencial de crear un nuevo “cinturón urbano” en la Sabana de Occidente, lo que fomentará la expansión inmobiliaria en zonas estratégicas y podrá atraer a empresas constructoras interesadas en proyectos residenciales y comerciales a largo plazo.

Sin embargo, el éxito de este impacto depende en gran medida de la correcta implementación y gestión del sistema ferroviario. Por esta razón, para maximizar los beneficios, es fundamental que el Regiotram opere con alta frecuencia y eficiencia, ya que una conectividad deficiente podría limitar su impacto en el mercado inmobiliario. Además, se espera que el proyecto incentive el desarrollo de infraestructuras de apoyo, como centros comerciales y espacios recreativos, alrededor de las estaciones, lo que incrementaría el valor de la zona y podría consolidar la Sabana como una extensión del área metropolitana de Bogotá.

Beneficios y Retos de la Infraestructura Vial en Cundinamarca

La infraestructura vial como el Regiotram aporta beneficios como la reducción de tiempos de viaje y una mejora en la accesibilidad para los habitantes de la Sabana de Occidente. Es importante destacar que este proyecto podría reducir el tiempo de desplazamiento entre Bogotá y municipios como Facatativá a menos de 55 minutos, lo cual es un cambio sustancial comparado con el promedio actual de hasta dos horas en condiciones de tráfico pesado (Castillo,

2024). Estos beneficios se reflejan en una mejor calidad de vida para los residentes y una mayor eficiencia en la movilidad urbana.

No obstante, el proyecto enfrenta retos importantes en términos de planificación urbana y sostenibilidad. Los municipios beneficiados tendrán que implementar regulaciones que permitan un desarrollo ordenado y que eviten problemas como la sobrepoblación o la saturación de servicios básicos. Asimismo, es de notar el riesgo de desplazamiento de residentes de bajos ingresos debido al incremento en el costo de vida. Por ello, es fundamental que la expansión de infraestructura como el Regiotram esté acompañada de políticas de inclusión social y de conservación ambiental, para asegurar que los beneficios económicos se distribuyan equitativamente.

Efectos sobre la Conectividad y Reducción de Tiempos de Desplazamiento

La conectividad que proporcionará el Regiotram representa uno de los efectos más positivos de este proyecto, con el potencial de reducir drásticamente los tiempos de desplazamiento y descongestionar las vías de acceso principales a Bogotá. Actualmente, el flujo de tráfico desde la Sabana a la capital está altamente saturado, y los viajes suelen demorar en promedio más de dos horas, especialmente en horas pico. Con el Regiotram, se estima que el tiempo de viaje hacia Bogotá se reducirá a menos de una hora, ofreciendo una alternativa rápida y directa para los habitantes de los municipios de Cundinamarca.

Este aumento en la conectividad no solo favorece a los residentes locales, sino que también crea nuevas oportunidades para quienes laboran en Bogotá, permitiendo que más personas puedan vivir en áreas suburbanas sin sacrificar un tiempo considerable en sus desplazamientos diarios. La mejora en la movilidad diaria incrementa el atractivo de estas áreas para residentes y empresas, incentivando la inversión en desarrollos habitacionales y comerciales

que satisfagan la creciente demanda. La conectividad que proporcionaría el Regiotram constituye un pilar fundamental para el crecimiento económico de la región y para la consolidación de la Sabana de Occidente como un polo de desarrollo.

Riesgos de Burbujas Inmobiliarias y Planificación Urbana Sostenible

Un riesgo destacado en este estudio monográfico es el de una posible burbuja inmobiliaria en los municipios aledaños al Regiotram, derivada de una valorización excesiva de los precios de las propiedades impulsada por la expectativa de una mejora en la conectividad. Tal como señala Gómez, (2023), una rápida valorización puede llevar a una sobrevaloración de propiedades y, eventualmente, a una corrección drástica en los precios, afectando tanto a inversionistas como a residentes locales. El aumento de precios debido a la especulación puede también desplazar a residentes de menores ingresos, limitando su acceso a zonas con mejor infraestructura y aumentando la desigualdad socioeconómica.

El documento enfatiza la necesidad de políticas de planificación urbana sostenible para mitigar estos riesgos. Medidas como la implementación de zonas de vivienda asequible, la regulación de precios y el control sobre el uso del suelo son esenciales para evitar un crecimiento desmedido e insostenible. Además, la planificación debe integrar espacios verdes y servicios públicos que aseguren una buena calidad de vida en las áreas urbanas emergentes. En resumen, este estudio monográfico resalta que la construcción del Regiotram debe ir acompañada de una estrategia de desarrollo urbano bien fundamentada, que promueva un crecimiento equilibrado y evite los peligros de una burbuja inmobiliaria.

Conclusiones

El análisis llevado a cabo en este trabajo resalta el impacto significativo que los proyectos de infraestructura ferroviaria similares al Regiotram de Occidente, pueden tener sobre los precios de los bienes inmuebles en las áreas cercanas a sus estaciones. A través de la revisión de estudios previos realizados en otras ciudades alrededor del mundo, se ha demostrado que la proximidad a estaciones de tren y la mejora en la conectividad entre ciudades o regiones son factores determinantes que influyen directamente en la valorización de las propiedades. Esto se observa en estudios realizados en ciudades como Phoenix, Teherán, Medellín y Madrid, donde se ha evidenciado un aumento en los precios de las viviendas cercanas a estaciones ferroviarias. En algunos casos, se ha registrado un incremento de entre el 10% y el 20%, lo que respalda la hipótesis de que el Regiotram de Occidente podría generar un impacto similar en la Sabana de Occidente.

Uno de los aspectos clave que emerge de este estudio es el rol fundamental que juega la mejora de la conectividad en el aumento del valor inmobiliario. La reducción significativa en los tiempos de desplazamiento que ofrecerá el Regiotram, al conectar municipios de Cundinamarca con Bogotá en menos de una hora, representa una mejora sustancial en la calidad de vida de los residentes. Este tipo de proyectos de infraestructura no solo facilita el acceso al centro urbano, sino que también incentiva la creación de nuevos polos de desarrollo, tanto residencial como comercial, en zonas previamente periféricas. El aumento en la demanda de propiedades cercanas a las estaciones del Regiotram puede generar un ciclo positivo de valorización en estas áreas, promoviendo un crecimiento económico que atraerá tanto a inversionistas como a compradores interesados en acceder a nuevas opciones habitacionales.

Sin embargo, al igual que en otros estudios internacionales sobre el impacto de proyectos de infraestructura ferroviaria, el análisis realizado también identifica ciertos riesgos asociados a este tipo de valorización. Uno de los principales peligros es el de la especulación inmobiliaria, que puede llevar a un aumento artificial de los precios de las propiedades debido a las expectativas generadas por el proyecto. Si bien estos aumentos iniciales en los precios pueden beneficiar a los primeros propietarios e inversionistas, podrían crear una burbuja inmobiliaria que afecte negativamente a los compradores potenciales, especialmente aquellos de estratos socioeconómicos más bajos, quienes podrían verse desplazados por los altos costos de vivienda. Este fenómeno podría generar exclusión social y aumentar la desigualdad en las zonas aledañas a las estaciones del Regiotram, lo que plantea un reto para las políticas urbanas y sociales en la región.

A pesar de los posibles riesgos, este estudio también subraya la importancia de implementar políticas públicas que promuevan un desarrollo urbano ordenado y sostenible. Para mitigar los efectos negativos de la especulación inmobiliaria, es esencial que los gobiernos locales y nacionales adopten medidas que promuevan la inclusión social, como la construcción de viviendas de interés social, la regulación de precios y el control del uso del suelo. Además, las autoridades deben trabajar para garantizar que los beneficios del Regiotram no se concentren únicamente en las zonas más privilegiadas, sino que se distribuyan de manera equitativa entre los diferentes estratos socioeconómicos. Esto es particularmente relevante en un contexto donde la presión sobre el mercado inmobiliario en la Sabana de Occidente está en aumento debido al crecimiento poblacional y la expansión urbana.

En conclusión, el Regiotram de Occidente representa una oportunidad única para transformar la región de la Sabana de Occidente, mejorando la conectividad y facilitando el

acceso a Bogotá y otras áreas clave. Si bien el impacto positivo sobre el mercado inmobiliario es evidente, los retos relacionados con la planificación urbana y la inclusión social deben ser considerados con detenimiento. Para que el proyecto sea verdaderamente exitoso, es crucial que se acompañe de una estrategia integral que contemple tanto el desarrollo de infraestructura como políticas de vivienda asequible, protección ambiental y participación ciudadana. Con una correcta implementación y gestión, el Regiotram tiene el potencial de replicar los beneficios observados en otros proyectos ferroviarios internacionales, consolidando a la Sabana de Occidente como un nuevo centro de desarrollo económico y urbano, sin perder de vista la necesidad de un crecimiento equilibrado y sostenible.

Recomendaciones

De acuerdo con los estudios analizados a lo largo de esta monografía, aunque el punto de partida común suele ser el desarrollo de modelos de machine learning clásicos, como las regresiones múltiples o los árboles de decisión, sería recomendable la implementación de un modelo híbrido. Este enfoque permitiría determinar de manera más precisa hasta qué punto un inversionista estaría dispuesto a pagar por una vivienda en el sector de la Sabana de Occidente. Con ello, se podrían evitar burbujas inmobiliarias que, como se mencionó previamente, tienden a inflar los precios basándose en especulaciones subjetivas, dificultando así el establecimiento de valores razonables y sostenibles.

La selección adecuada de variables es un factor determinante no solo para la aplicación del modelo, sino también para su éxito. En este sentido, se recomienda incluir no solo las variables intrínsecas y clásicas relacionadas con las características de la vivienda, sino también variables externas, especialmente si el objetivo es evaluar aspectos relacionados con la zona. En el contexto de este estudio, una variable clave identificada fue la cercanía a una estación del Regiotram de Occidente. Sin embargo, este principio también es aplicable a otros elementos de interés, como el impacto de parques, centros de salud, centros comerciales u otras infraestructuras relevantes. Así mismo, es importante integrar análisis temporales en el modelo híbrido, considerando la evolución de los precios de las viviendas en el tiempo, esto podría ayudar a identificar tendencias y patrones que permitan prever posibles fluctuaciones y a los inversionistas tomar decisiones más informadas.

Se recomienda incorporar herramientas de análisis geoespacial que permitan evaluar de forma más detallada la influencia de la ubicación en los precios de las viviendas. Esto incluye

factores como la accesibilidad al transporte, la densidad poblacional, y la proximidad a zonas de alto desarrollo económico.

Por otro lado, para garantizar la efectividad del modelo híbrido, es crucial realizar validaciones periódicas con nuevos conjuntos de datos. Esto permitirá ajustar los parámetros del modelo a las dinámicas cambiantes del mercado inmobiliario y asegurar la precisión en sus predicciones. También es recomendable involucrar a actores clave como urbanistas, agentes inmobiliarios y economistas en el desarrollo del modelo. Esto no solo enriquecerá la calidad de las variables seleccionadas, sino que también garantizará que el modelo sea funcional y adaptable a las necesidades reales del mercado.

Finalmente, los resultados del modelo podrían utilizarse para respaldar políticas públicas que regulen el mercado inmobiliario, fomentando precios más justos y evitando prácticas especulativas que afecten la sostenibilidad del sector.

Enlace a Diapositivas. Haga [click aquí](#)

Referencias Bibliográficas

- Abdulla, H., Ibrahim, M., & Al-Hinkawi, W. (2023). The Impact of Urban Street Network on Land Value: Correlate Syntactical Premises to the Land Price. *Buildings*, *13*(7), 1610. <https://doi.org/10.3390/buildings13071610>
- Abdul-Rahman, S., Mutalib, S., & Zulkifley, N. (2024). Advanced Machine Learning Algorithms for House Price Prediction: Case Study in Kuala Lumpur. *International Journal of Advanced Computer Science and Applications*, *12*(12), 736-745. <https://doi.org/10.14569/IJACSA.2021.0121291>
- Abidoye, R., Fam, F., Oshodi, O., & Oyetunji, A. (2021). Impact of light rail line on residential property values – a case of Sydney, Australia. *International Journal of Housing Markets and Analysis*, *15*(3), 691-708. <https://doi.org/10.1108/IJHMA-03-2021-0033>
- Alonzo, D. (2017). *El tren vuelve a la sabana: Reflexiones sobre los posibles impactos del Regiotram en sectores urbanos de los municipios de Funza, Mosquera y Madrid, en el marco del proceso de metropolización* [Tesis de Maestría, Pontificia Universidad Javeriana]. <http://repository.javeriana.edu.co/handle/10554/40647>
- Ayuso, J., & Restoy, F. (2006). House prices and rents: An equilibrium asset pricing approach. *Journal of Empirical Finance*, *13*(3), 371-388. <https://doi.org/10.1016/j.jempfin.2005.10.004>
- Castillo, C. (2024, 15 de junio). *Pronto iniciarán las obras del Regiotram de Occidente: Avances claves para conectar a Bogotá con Facatativá en minutos*. Red+ Noticias. <https://redmas.com.co/colombia/Pronto-iniciaran-las-obras-del-Regiotram-de-Occidente-avances-claves-para-conectar-a-Bogota-con-Facatativa-en-minutos-20240915-0013.html>

- Debrezion, G., Pels, E., & Rietveld, P. (2007). The Impact of Railway Stations on Residential and Commercial Property Value: A Meta-analysis. *The Journal of Real Estate Finance and Economics*, 35(2), 161-180. <https://doi.org/10.1007/s11146-007-9032-z>
- Escobar, A. (2019, 16 de mayo). *5 Factores que inciden en la valorización de un inmueble*. Camacol Antioquia. <https://camacolantioquia.org.co/5-factores-que-inciden-en-la-valorizacion-de-un-inmueble/>
- Forouhar, A. (2016). Estimating the impact of metro rail stations on residential property values: Evidence from Tehran. *Public Transport*, 8(3), 427-451. <https://doi.org/10.1007/s12469-016-0144-9>
- García, P. (2021). *Implementación de un modelo machine learning para la estimación del valor del metro cuadrado de un inmueble ubicado en Cundinamarca* [Tesis de Maestría]. Universidad de los Andes.
- Gil, J., Toro, G., & Villante, J. (2022). *Guía De Estructuración De Proyectos De Infraestructura De Transporte. Dirección Técnica Y De Estructuración Subdirección De Estructuración De Proyectos*. Instituto Nacional de Vías.
- Gómez, N. (2023,25 de septiembre). *¿Qué es una burbuja inmobiliaria y cómo puede afectar en la economía colombiana?* El Tiempo. <https://www.eltiempo.com/economia/sector-financiero/que-es-una-burbuja-inmobiliaria-y-como-afecta-la-economia-de-colombia-818923>
- González, F., Guzmán, T., & Barrera, V. (2015). *Estrategias para la construcción de paz territorial en Colombia. Elementos para la discusión*. Cinep.
- Grajales, Y. (2019). *Modelo de predicción de precios de viviendas en el Municipio de Rionegro para apoyar la toma de decisiones de compra y venta de propiedad raíz* [Tesis de

Maestría, Universidad Pontificia Bolivariana].

<https://repository.upb.edu.co/handle/20.500.11912/5285>

Lederman, J. (2023, 17 de enero). *Especular en inversiones: Bienes Raíces*. Finanzas y Riqueza.

<https://es.linkedin.com/pulse/especular-en-inversiones-bienes-ra%C3%ADces-jonathan-lederman>

Ministerio de Transporte. (2024, 30 de octubre). *Regiotram de occidente*. Proyectos Ferreos.

<https://mintransporte.gov.co/publicaciones/11651/regiotram-de-occidente/>

Munevar, O. (2016). *La estadística en valoración inmobiliaria y la aplicación de los avalúos en la gestión del suelo* [Tesis de Especialización, Universidad Santo Tomás].

<https://hdl.handle.net/11634/2304>

Özögür, S., Eygi, B., Yıldız, Ö., & Karadayı, P. (2023). A Novel Hybrid House Price Prediction

Model. *Computational Economics*, 62(3), 1215-1232. <https://doi.org/10.1007/s10614-022-10298-8>

ProBogotá. (2022). *Plan Estratégico de Desarrollo para Sabana Occidente a 2051*. Fundación para el progreso de la Región.

Puentes, N., & Garza, J. (2023). Radical uncertainty and the effect of transport infrastructure on

land prices. *Brazilian Journal of Political Economy*, 43(3), 747-769. <https://doi-org.bibliotecavirtual.unad.edu.co/10.1590/0101-31572023-3374>

Rodríguez, H., Carrero, D., & García, W. (2021). *Dinámica inmobiliaria Bogotá—Región*

2017—2020: Un estudio de la oferta de vivienda usada en Bogotá y la región de Cundinamarca en el periodo 2017-2020. Observatorio Técnico Catastral.

Selim, H. (2009). Determinants of house prices in Turkey: Hedonic regression versus artificial neural network. *Expert Systems with Applications*, 36(2), 2843-2852.

<https://doi.org/10.1016/j.eswa.2008.01.044>

Seo, K., Salon, D., Kuby, M., & Golub, A. (2019). Hedonic modeling of commercial property values: Distance decay from the links and nodes of rail and highway infrastructure. *Transportation*, 46, 859-882.

Tyagi, Y., & Singhal, S. (2018). Analyzing the Influence of Metro Rail Ridership on Group Housing Property Prices in Delhi. *Real Estate Finance Journal*, 35(1), 10-22.