



Propuesta Proyecto

Tasación automática de propiedades

Grupo 23 — Javier Cuitiño, Simón Saravia, Alejandro Orellana

21 de octubre de 2025

Contexto y motivación

Pregunta: Discuta la motivación y razones para su elección del proyecto, incluyendo el contexto y alguna investigación que haya incluido en su decisión. Comente también sobre la audiencia objetivo de su análisis (quién será el tomador de decisiones que se apoyará en su trabajo).

Respuesta: La compra-venta de viviendas y departamentos exige que ambas partes (comprador y vendedor) cuenten con un valor de referencia objetivo. Una metodología ampliamente utilizada es la valoración de mercado basada en propiedades publicadas en portales inmobiliarios. El proyecto busca automatizar esa tasación aprovechando la disponibilidad de información pública.

Preguntas objetivo

Pregunta: ¿Cuáles son las preguntas que busca responder mediante su análisis?

Respuesta: En este caso, responderemos las preguntas de investigación que habíamos detallado en nuestra propuesta.

1. ¿Cuáles son los bienes más similares (comparables) a la propiedad en estudio?

a) No se puede responder aún.

2. ¿Es posible construir una muestra representativa de comparables?

a) Sí, es posible construir una muestra representativa de comparables si es que tenemos en cuenta una correcta separación de los datos. Como se puede observar en los gráficos, la dispersión de los precios principalmente es muy alta, debido a que una propiedad no solo depende de sus características para su precio, si no también de su ubicación.

Para construir correctamente una muestra representativa de comparables es primordial separar correctamente los datos por comunas para poder comparar todo de manera correcta y que no se vean afectadas las muestras representativas por el alto nivel de dispersión de los datos.

3. ¿Cuál es el nivel de dispersión de los precios en la muestra?
 - a) El nivel de dispersión de los precios en la muestra es extremadamente alto. Como podemos observar en los gráficos, la dispersión se ve como un amplio barrido de puntos en el eje vertical, principalmente debido a que estamos comparando muchas propiedades con distintas características o de diferentes sectores, lo que es el principal motivo de las diferencias de precio.
4. ¿Es posible aproximar/identificar ventas reales frente a simples publicaciones?
 - a) No se puede responder aún.
5. ¿Cuál ha sido la evolución histórica de los precios en el área?
 - a) Debido a la complejidad para acceder a los datos, no pudimos encontrar datos con respecto a la evolución histórica de los precios del área. Solamente encontramos datos actuales, por lo que no fue posible establecer una relación con el cambio de precios durante el tiempo.

Datos

Pregunta: Describa los datos utilizados y su origen, y documente los procedimientos de recolección, preparación y transformación de datos. Describa los datos indicando el tipo de datos, su estructura y los elementos que se describen en la base de datos.

Respuesta: Los datos trabajados se obtuvieron mediante web scraping, inicialmente se sacaron dos datasets, uno de casas y otro de departamentos, cada uno con distintas columnas importantes para luego proceder a hacer la regresión y poder obtener precios a partir de inputs, los dos datasets se juntaron y se añadió la columna "tipo de vivienda", luego había olvidado un dato importante a analizar que es la dirección de cada casa, así que se volvió a hacer scraping esta vez llenando url por url sacando solo la dirección de cada vivienda y finalmente concatenando ambos datasets en uno final. Luego se trabajó con el dataset final, se sacaron filas con valores nulos y se limpiaron outliers.

Análisis exploratorio de datos

Pregunta: Describa el proceso de exploración de sus datos, incluyendo visualizaciones, análisis gráficos, análisis estadísticos, etc. Explique los resultados y conclusiones preliminares obtenidas, y cómo estos influyen o motivan el método de modelamiento elegido. Justifique todas las decisiones adoptadas, tanto la elección de herramientas como las elecciones de visualización.

Respuesta: La respuesta está en el notebook del repositorio.

Link repositorio: <https://github.com/Saaay20/PROYECTOIMT2200>