

Taller 1

Reyes Gómez Maria Paola
{mp.reyesg1}@uniandes.edu.co
Instructor: Fabian Camilo Peña
MINE4101
22 de Septiembre 2024

El objetivo de este taller es evaluar las capacidades del estudiante para realizar análisis de datos utilizando técnicas estadísticas y de visualización, que le permitan entender un conjunto de datos, descubrir insights y sugerir accionables al negocio.

Contexto de negocio: Inversiones Inmobiliarias para alquiler vacacional.

Como consultor científico de datos especializado en el sector inmobiliario, su objetivo es ayudar a individuos o empresas que buscan invertir en propiedades para alquiler a través de la plataforma Airbnb. Con la creciente popularidad de este modelo de negocio, es cada vez más difícil tomar buenas decisiones respecto a los mejores sectores de la ciudad, tipos de propiedad, entre otros factores, para lograr una rentabilidad adecuada a lo largo del tiempo. Una vez que la propiedad se encuentra listada en la plataforma, asegurar una alta tasa de ocupación estableciendo un precio por noche justo y un buen sistema de atención a los usuarios es vital para el éxito del negocio.

El dataset de trabajo utilizado es el de la ciudad de Bruselas, Bélgica.

I. ENTENDIMIENTO INICIAL DE LOS DATOS

I-A. Dimensiones del dataset:

El dataset contiene 6119 registros y 75 columnas.

I-B. Tipo de Datos:

La mayoría de columnas contienen datos de tipo object (cadenas de texto), mientras que algunas contienen valores numéricos (int64, float64).

I-C. Top 5 de atributos más importantes:

Para el análisis de inversión inmobiliaria, he identificado 5 atributos clave:

- **price:** Precio de la propiedad, esencial para evaluar el valor de las inversiones. El precio no aparece directamente en el resumen debido a un posible formato de texto, pero se requiere análisis detallado.
- **neighbourhood_cleansed:** Zona donde se encuentra la propiedad, importante para analizar la ubicación y rentabilidad potencial.
- **number_of_reviews:** Cantidad de reseñas de cada propiedad, indica popularidad.
- **review_scores_rating:** Calificación promedio de la propiedad, refleja la calidad percibida.
- **bedrooms:** Número de habitaciones, esencial para categorizar tipos de propiedades.

I-D. Análisis Multivariado preliminar

Para el análisis multivariado de la inversión inmobiliario quedó así:

- **price:** Se necesita transformar este campo para obtener un análisis de distribución.
- **neighbourhood_cleansed:** Se distribuye en varias categorías, representando las diferentes zonas.
- **number_of_reviews:** Tiene una media de 23.43 reseñas, con un máximo de 629, lo que indica una alta variabilidad.
- **review_scores_rating:** El valor promedio es 4.67, lo que sugiere que las propiedades tienden a tener buenas calificaciones.
- **bedrooms:** La mayoría de las propiedades tiene entre 1 y 2 habitaciones, con un máximo de 8.

II. ESTRATEGIA DE ANÁLISIS

Para determinar la mejor opción de inversión inmobiliaria, sugiero una estrategia basada en el análisis detallado de varios atributos clave: precio, ubicación, número de reseñas, calificación de la propiedad y características físicas (número de habitaciones y baños). Primero comienzo con un análisis descriptivo calculando estadísticos básicos como la media, mediana, rango y percentiles para cada variable, lo que me permitirá identificar propiedades con características promedio o fuera de lo común. Esto incluye evaluar las distribuciones de precios y calificaciones por vecindario, buscando áreas que presenten una relación favorable entre precio y calidad percibida. También identificaremos vecindarios con alta demanda analizando el número de reseñas y la ocupación estimada a través de métricas como `reviews_per_month`.

A nivel gráfico, utilizaré técnicas como mapas de calor geográficos para visualizar la concentración y rango de precios y calificaciones en diferentes zonas de la ciudad, lo que ayudará a identificar áreas de mayor rentabilidad. Además, gráficos de dispersión cruzarán atributos como precio versus calificación o precio versus número de reseñas, para visualizar tendencias y correlaciones. Las distribuciones de densidad y boxplots permitirán detectar outliers o propiedades con características inusuales, que pueden representar oportunidades de inversión atípicas. Finalmente, el uso de gráficos multivariados como "pair plots" o diagramas de burbuja permitirá analizar múltiples factores simultáneamente y tomar decisiones más informadas al recomendar áreas y tipos de propiedades que ofrecen un equilibrio óptimo entre costo, demanda y calidad.

III. DESARROLLO DE LA ESTRATEGIA

1. **Procesamiento de Datos:** Primero, se limpia y transforman algunos atributos clave, como el precio y la ubicación, para hacerlos adecuados para el análisis. Luego, se imputa valores faltantes, y se convierten algunas columnas a tipos de datos numéricos donde sea necesario. Además, se puede necesitar manejar atributos categóricos como el vecindario o los tipos de propiedad.
2. **Cálculo de Estadísticos Básicos:** Usaré estadísticas descriptivas (media, mediana, percentiles) para obtener una visión general de la distribución de los precios, calificaciones, número de reseñas, y otras características importantes de las propiedades.
3. **Visualización de Datos:** A través de diversas técnicas gráficas, podremos comprender la relación entre variables como precio, ubicación, reseñas y calificación. Para ello usaremos Mapas de calor para visualizar la distribución de los precios y calificaciones por vecindario. Gráficos de dispersión para detectar la correlación entre precio y número de reseñas, y entre precio y calificación. Boxplots y Distribuciones de rango y precio para identificar outliers y patrones de variabilidad dentro de cada vecindario y rango de precio.

IV. GENERACIÓN DE RESULTADOS

Informe Ejecutivo: Recomendaciones de Inversión Inmobiliaria Basada en Datos

1. **Análisis de Mercado** El análisis de datos de las propiedades muestra una amplia variabilidad en precios y características, dependiendo de su ubicación y atributos. A partir de los datos, se identifican patrones de demanda y características que permiten seleccionar las mejores oportunidades de inversión.
2. **Sectores de Inversión** Basado en el análisis de precios promedio por vecindario, se destacan los siguientes sectores como atractivos para invertir como vecindarios más caros, vecindarios más asequibles, entre otros.

Para el caso del análisis de precios promedio del vecindario más caro, estas áreas ofrecen propiedades con precios promedio más altos en donde el mercado orientado al segmento del lujo o de alta demanda. Las propiedades en estos sectores suelen atraer a inversores interesados en maximizar ingresos por alquiler a corto y largo plazo. Por otro lado, los vecindarios más asequibles las áreas donde el precio es más acorde con potencial de apreciación y alta demanda, lo anterior obedece a que precios más bajos implican una alta ocupación y al final una buena reseña y con retornos estables con menores costos iniciales.

Los tipos de propiedades normalmente muestra en el análisis de las propiedades con 2 a 3 habitaciones son las que tienen mayor demanda en zonas urbanas o turísticas. Los tres aspectos más importantes a tener en cuenta es el número de habitaciones, vecindarios turísticos y comerciales, cazas o town así como las zonas familiares.

Para las calificaciones y reseñas, en términos generales, las propiedades con mejores calificaciones y reseñas son las que suelen ser las mejor mantenidas y ofrecer una experiencia que es satisfactoria para el cliente. Los enfoques que maximizan estas características son el mantenimiento y la limpieza y el número de reseñas.

Adicionalmente, los rangos de precios dependiente de las propiedades con mayor éxito y demanda se encuentra dentro de unos rangos de precio establecidos y ofrecen una combinación ideal entre el precio, demanda, calidad y con el número de reseñas positivas y las calificaciones altas, asegurando la alta ocupación.

Para los puntos clave los servicios adicionales son los factores clave en las calificaciones y ocupación de las propiedades. Aspectos como WIFI gratuito, aire acondicionado o calefacción, espacio de trabajo y acceso a otras zonas fuera del sitio son factores que suman mucho.

Para el análisis de riesgo aunque los vecindarios premium tienen un mayor potencial de retorno, también hay un riesgo que se asocia con tiempos de vacaciones largos y de mayor sensibilidad a las fluctuaciones del mercado inmobiliario. Parte de la recomendación es diversificar la mayoría de inversiones entre propiedades premium y más asequibles equilibrando tanto las zonas y los tipos de inquilinos.

Basado en los datos, se recomienda enfocar la inversión en sectores para propiedades premium, algunos vecindarios particulares ofrecen oportunidades más asequibles con alta demanda. Las propiedades de 2 a 3 habitaciones ofrecen el mejor equilibrio entre costo y retorno, especialmente si se optimizan con servicios adicionales valorados por los usuarios.