

Análisis Exploratorio de los Datos de Propiedades del portal Remax

1. Descripción General de los Datos

El dataset utilizado para el análisis contiene 48 registros de propiedades listadas en la ciudad de Quito, Ecuador, por la plataforma inmobiliaria Remax. Los datos incluyen diversas características de las propiedades como el precio, las expensas, la dirección, la superficie total y cubierta, el número de ambientes y baños, así como descripciones textuales que brindan una visión de las propiedades. También se incluye información del agente inmobiliario encargado de la venta y el enlace directo a la propiedad.

La estructura de las columnas es la siguiente:

- **Precio:** Monto total de la propiedad en dólares estadounidenses (USD).
- **Expensas:** Gastos mensuales adicionales (en USD), como mantenimiento o administración.
- **Dirección:** Ubicación exacta de la propiedad.
- **Superficie Total (m²):** Tamaño total de la propiedad.
- **Superficie Cubierta (m²):** Área construida o techada de la propiedad.
- **Ambientes:** Número de ambientes que posee la propiedad.
- **Baños:** Número total de baños.
- **Descripción:** Texto descriptivo que detalla características adicionales de la propiedad.
- **Agente:** Nombre del agente encargado de la propiedad.
- **URL Propiedad:** Enlace al listado de la propiedad en la web.

Este conjunto de datos ofrece una buena variedad de información que permite realizar un análisis profundo del mercado inmobiliario en Quito y los diferentes factores que pueden influir en el precio de las propiedades.

2. Análisis de Tipos de Datos

Los tipos de datos en este conjunto se dividen en dos principales categorías: numéricos y categóricos (de tipo texto). A continuación se describen las principales columnas:

1. Numéricos:

- Precio: int64
- Expensas: float64
- Superficie Total (m²): int64
- Superficie Cubierta (m²): int64
- Ambientes: int64
- Baños: int64

2. Categóricos (Texto):

- Dirección: object
- Descripción: object
- Agente: object

- URL Propiedad: object

Este desglose nos permite observar que las columnas con datos numéricos están correctamente formateadas, lo cual es crucial para realizar análisis cuantitativos, como la correlación entre variables o el clustering de propiedades.

3. Análisis Exploratorio de las Variables Numéricas

El análisis de las variables numéricas es crucial para comprender mejor cómo se distribuyen los datos y qué relaciones pueden existir entre ellas. En este dataset de Remax, nos enfocamos en dos variables principales: el **precio de las propiedades** y su **superficie** (tanto total como cubierta). Estas variables proporcionan una visión clara de cómo se comporta el mercado inmobiliario en Quito.

Distribución de Precios

El precio de una propiedad es uno de los factores más importantes en la compra y venta de inmuebles. Como se observa en el gráfico de la **distribución de precios**, la mayoría de las propiedades se encuentran en el rango de \$50,000 a \$150,000 USD, lo que sugiere que este segmento del mercado está dirigido principalmente a compradores de clase media. Además, se puede ver que existen pocas propiedades en el rango de lujo, con precios que superan los \$300,000 USD.

Relación entre Precio y Superficie

La relación entre el precio de una propiedad y su superficie total es clave para determinar si el tamaño de la propiedad está directamente vinculado con su precio. El gráfico de dispersión que muestra la **relación entre superficie y precio** revela que, como era de esperar, las propiedades más grandes tienden a ser más caras. Sin embargo, hay algunas propiedades con superficies relativamente pequeñas pero precios elevados, lo que podría deberse a su ubicación o características especiales, como vistas panorámicas o amenities exclusivas.

La **superficie cubierta** también muestra una correlación positiva con el precio, aunque hay algunas excepciones. Esto sugiere que las propiedades con más área construida tienden a ser más valiosas, pero otros factores, como la ubicación, también pueden influir significativamente en el precio final.

Relación entre Gastos y Precio

El análisis de la **relación entre gastos y precio** revela que las propiedades con mayores gastos tienden a estar en el rango de precios más altos. Sin embargo, no se trata de una relación perfecta, ya que algunas propiedades de lujo tienen gastos relativamente bajos, lo que podría indicar diferencias en la administración o en los servicios incluidos en los gastos. Las propiedades con gastos más bajos se agrupan en el rango de \$50,000 a \$150,000 USD, lo cual es consistente con las propiedades de tamaño mediano que dominan el mercado.

Distribución de Superficies

El análisis de la **distribución de la superficie total** y la **superficie cubierta** de las propiedades revela que la mayoría de las propiedades tienen menos de 5000 m² de superficie total, y entre 50 y 200 m² de superficie cubierta. Esto refleja un mercado orientado a propiedades de tamaño medio, probablemente apartamentos o casas urbanas. Sin

embargo, hay algunas propiedades de gran tamaño, posiblemente ubicadas en áreas más exclusivas o rurales, que superan los 15,000 m².

Relación entre Precio y Número de Ambientes/Baños

El análisis de la relación entre el número de **ambientes y baños** y el precio de la propiedad sugiere que las propiedades con más ambientes y baños tienden a ser más costosas, lo cual es esperable. Los boxplots revelan una mayor dispersión de precios en propiedades con más de 5 ambientes o más de 3 baños, lo que indica que estas propiedades pueden variar ampliamente en función de otros factores, como la ubicación o las características de lujo.

4. Evaluación Inicial de Datos Faltantes

En la evaluación de los datos, se encontraron valores faltantes en la columna de **expensas**, donde algunas propiedades no incluyen este dato. Esto puede deberse a que ciertas propiedades no están asociadas a gastos comunes o que la información simplemente no fue proporcionada. Dado que solo el 17% de los registros tienen datos faltantes en esta columna, y que las expensas no son un factor crítico para todos los análisis, se puede optar por dejar esta columna tal como está o realizar imputaciones con promedios según el rango de precios.

El resto de las columnas numéricas, como **precio, superficie y ambientes**, están completas, lo que asegura la solidez de los análisis realizados. Las columnas de **descripción y agente** también están completas, lo que facilita la comprensión de las características adicionales de las propiedades y el análisis de palabras clave. En futuras iteraciones, podría ser útil investigar si la inclusión de más datos sobre amenities o características específicas (como parqueaderos, piscinas, etc.) puede enriquecer el análisis.

En resumen, el análisis de las variables numéricas y la evaluación de los datos faltantes muestra que el dataset de Remax es robusto y ofrece una buena base para comprender el comportamiento del mercado inmobiliario en Quito. A pesar de algunos valores faltantes, la calidad general de los datos es adecuada para análisis cuantitativos y visuales, lo que permite extraer conclusiones significativas sobre el mercado.

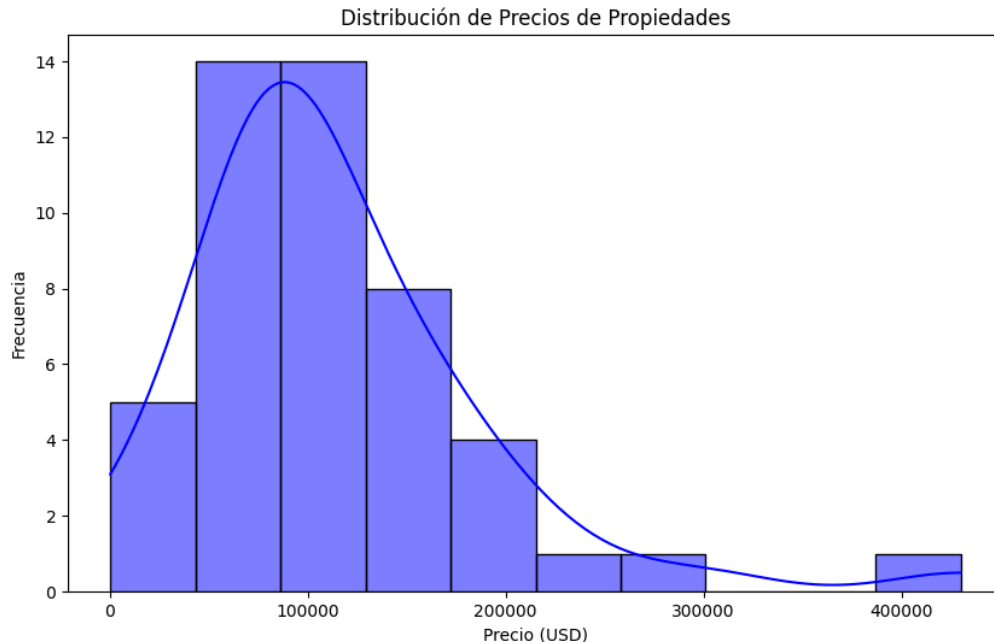
Análisis Visual de las Variables Clave

A continuación, se generarán gráficos que permitirán explorar la distribución de precios, la relación entre variables, y la existencia de valores atípicos.

1. Distribución de Precios

```
plt.figure(figsize=(10, 6))
sns.histplot(df_remax['Precio'], bins=10, kde=True, color='blue')
plt.title('Distribución de Precios de Propiedades')
plt.xlabel('Precio (USD)')
plt.ylabel('Frecuencia')
```

```
plt.savefig('images/Remax/Remax_exp_precio_dist.png')
plt.show()
```



Descripción: Este histograma muestra la distribución de precios de las propiedades listadas en el mercado de Remax. Nos permite observar la frecuencia de diferentes rangos de precios en el conjunto de datos.

Observación General: La mayoría de las propiedades están en el rango de \$50,000 a \$150,000. Sin embargo, existen algunas propiedades de lujo cuyo precio supera los \$300,000, aunque estas son mucho menos frecuentes.

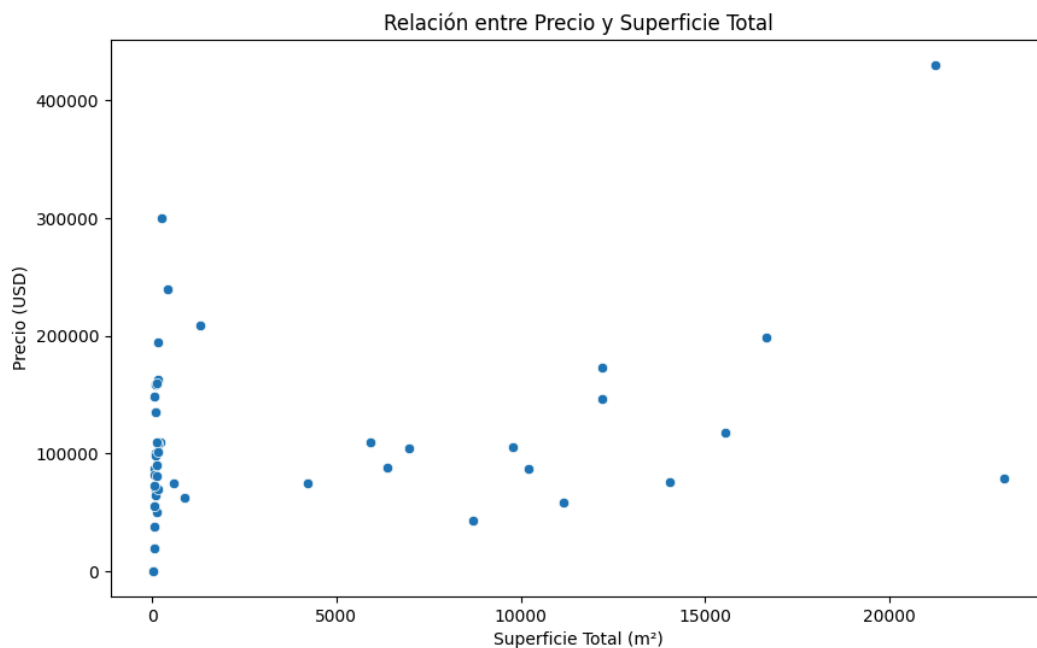
Conclusiones:

- **Mercado Enfocado en Precios Moderados:** El mercado de Remax parece estar mayormente enfocado en propiedades de precio moderado, con la mayoría de los inmuebles concentrados en el rango de \$50,000 a \$150,000. Esto sugiere que el público objetivo son familias de clase media o compradores primerizos.
- **Poca Oferta de Propiedades de Lujo:** Propiedades de más de \$300,000 son una minoría en este mercado, lo que sugiere que el segmento de lujo está menos representado en Remax.

Recomendación: Aumentar la visibilidad de las propiedades de lujo en campañas dirigidas a inversionistas o compradores de alto poder adquisitivo. Al mismo tiempo, para las propiedades más accesibles, enfocarse en resaltar las ventajas de la relación calidad-precio en áreas urbanas y suburbanas.

2. Relación entre Precio y Área

```
plt.figure(figsize=(10, 6))
sns.scatterplot(x='Superficie Total (m²)', y='Precio',
data=df_remax)
plt.title('Relación entre Precio y Superficie Total')
plt.xlabel('Superficie Total (m²)')
plt.ylabel('Precio (USD)')
plt.savefig('images/Remax/Remax_exp_precio_vs_superficie.png')
plt.show()
```



Descripción: Este gráfico de dispersión muestra la relación entre el precio de las propiedades y su superficie total, con el objetivo de identificar una correlación entre estas dos variables.

Observación General: Aunque se observa una tendencia a que propiedades más grandes tengan precios más elevados, también hay propiedades con grandes superficies a precios relativamente bajos. Esto sugiere que el precio no solo depende del tamaño de la propiedad, sino también de otros factores como la ubicación o el estado de la construcción.

Conclusiones:

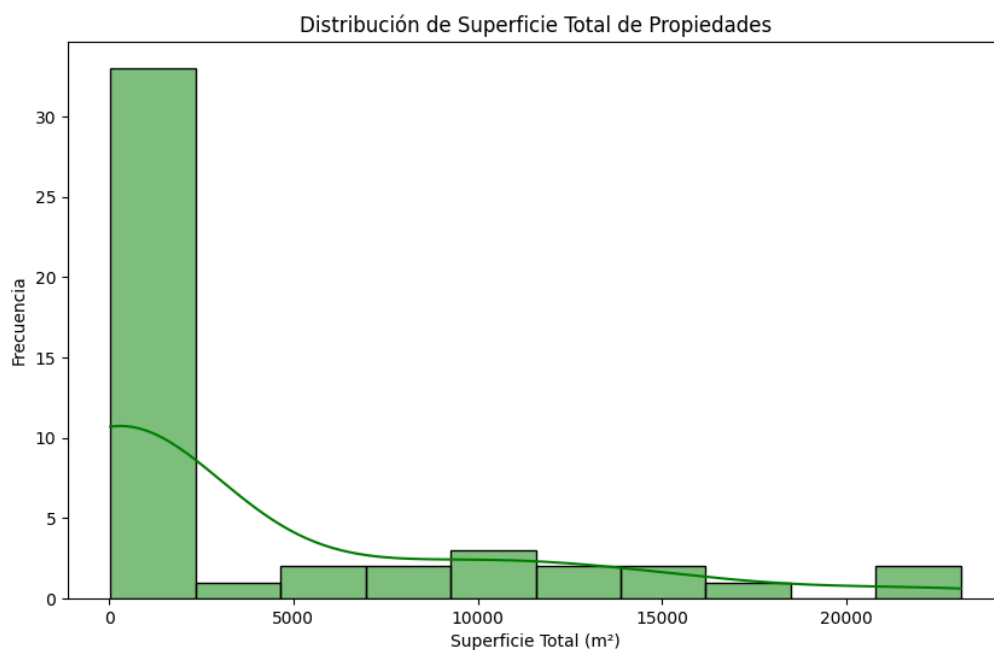
- **Relación Precio-Tamaño:** En general, las propiedades con superficies mayores tienen precios más altos, lo que es esperable en el mercado inmobiliario. Sin embargo, se observan algunas propiedades de gran tamaño con precios menores a \$100,000, lo que podría indicar una menor demanda o ubicación en zonas periféricas.

- **Propiedades Pequeñas y Caras:** Algunas propiedades con superficies pequeñas tienen precios elevados, lo que podría deberse a su ubicación en áreas más exclusivas o al estado de las construcciones.

Recomendación: Promover las propiedades grandes pero asequibles como oportunidades de inversión para aquellos que buscan terrenos o casas amplias fuera del centro urbano. Las propiedades pequeñas pero costosas pueden dirigirse a compradores que valoran más la ubicación que el tamaño.

3. Distribución de Áreas

```
plt.figure(figsize=(10, 6))
sns.histplot(df_remax['Superficie Total (m²)'], bins=10, kde=True,
color='green')
plt.title('Distribución de Superficie Total de Propiedades')
plt.xlabel('Superficie Total (m²)')
plt.ylabel('Frecuencia')
plt.savefig('images/Remax/Remax_exp_distribucion_superficie.png')
plt.show()
```



Descripción: Este histograma muestra la distribución de la superficie total de las propiedades listadas. Se observa cómo varía el tamaño de las propiedades en términos de metros cuadrados.

Observación General: La mayoría de las propiedades tienen superficies menores a 5,000 m², lo que indica que el mercado está compuesto principalmente por propiedades urbanas o suburbanas de tamaño moderado.

Conclusiones:

- **Predominio de Propiedades Pequeñas a Medianas:** La mayoría de las propiedades tienen una superficie total menor a 5,000 m². Esto es indicativo de un mercado inmobiliario enfocado en propiedades urbanas o suburbanas, donde las parcelas son más compactas.
- **Segmento de Propiedades Grandes:** Un pequeño porcentaje de propiedades tiene superficies significativamente mayores, superando los 15,000 m², lo que sugiere la existencia de un segmento de propiedades rurales o terrenos destinados a la construcción de proyectos más grandes.

Recomendación: Las propiedades más grandes podrían promocionarse como terrenos con potencial para proyectos inmobiliarios o comerciales. Sin embargo, la mayor parte del mercado parece estar más interesada en propiedades más accesibles en términos de tamaño, lo que debería influir en las estrategias de marketing.

4. Relación entre Precio y Gastos

```
plt.figure(figsize=(10, 6))
sns.scatterplot(x='Gastos', y='Precio', data=df_remax)
plt.title('Relación entre Gastos y Precio de Propiedades')
plt.xlabel('Gastos (USD)')
plt.ylabel('Precio (USD)')
plt.savefig('imagenes/Remax/Remax_exp_precio_vs_gastos.png')
plt.show()
```



Descripción: Este gráfico de dispersión muestra la relación entre el costo mensual de las expensas y el precio total de las propiedades.

Observación General: Existe una correlación positiva entre las expensas y el precio de las propiedades, lo que sugiere que propiedades más costosas tienden a tener expensas más altas, probablemente debido a la inclusión de más servicios y amenidades.

Conclusiones:

- **Expensas y Lujo:** Las propiedades con expensas más altas tienden a ser aquellas con precios elevados, lo que sugiere que están ubicadas en edificios con más servicios o en barrios más exclusivos.
- **Algunas Excepciones:** También se observan propiedades con expensas bajas y precios altos, lo que puede indicar que los servicios comunes no son un factor determinante en esas propiedades.

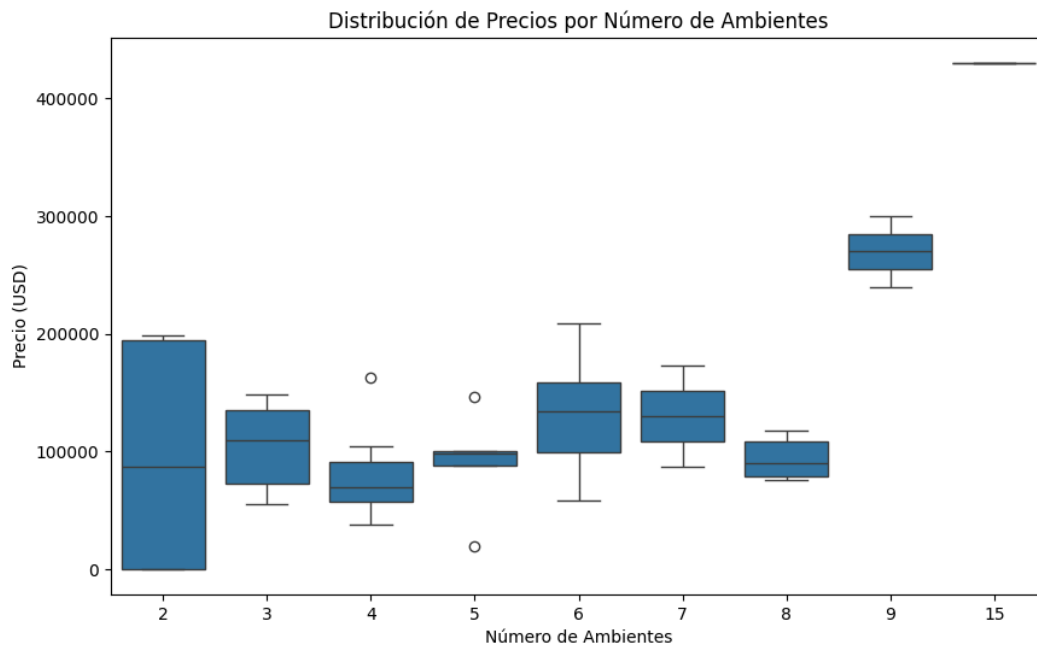
Recomendación: Evaluar la importancia de las expensas al comercializar propiedades. Propiedades con expensas bajas podrían ser promocionadas como una opción más accesible en términos de costos recurrentes.

5. Distribución de Precios por Número de Ambientes

```
plt.figure(figsize=(10, 6))
sns.boxplot(x='Ambientes', y='Precio', data=df_remax)
plt.title('Distribución de Precios por Número de Ambientes')
plt.xlabel('Número de Ambientes')
plt.ylabel('Precio (USD)')
```



```
plt.savefig('images/Remax/Remax_exp_precio_por_ambientes.png')  
plt.show()
```



Descripción: Este gráfico de caja muestra la relación entre el número de ambientes en una propiedad y su precio. Se busca identificar cómo varía el precio a medida que una propiedad ofrece más espacios habitables.

Observación General: A mayor número de ambientes, los precios tienden a incrementarse, aunque existe una amplia variabilidad en propiedades con más de 5 ambientes, lo que indica que otras variables como la ubicación también influyen en el precio.

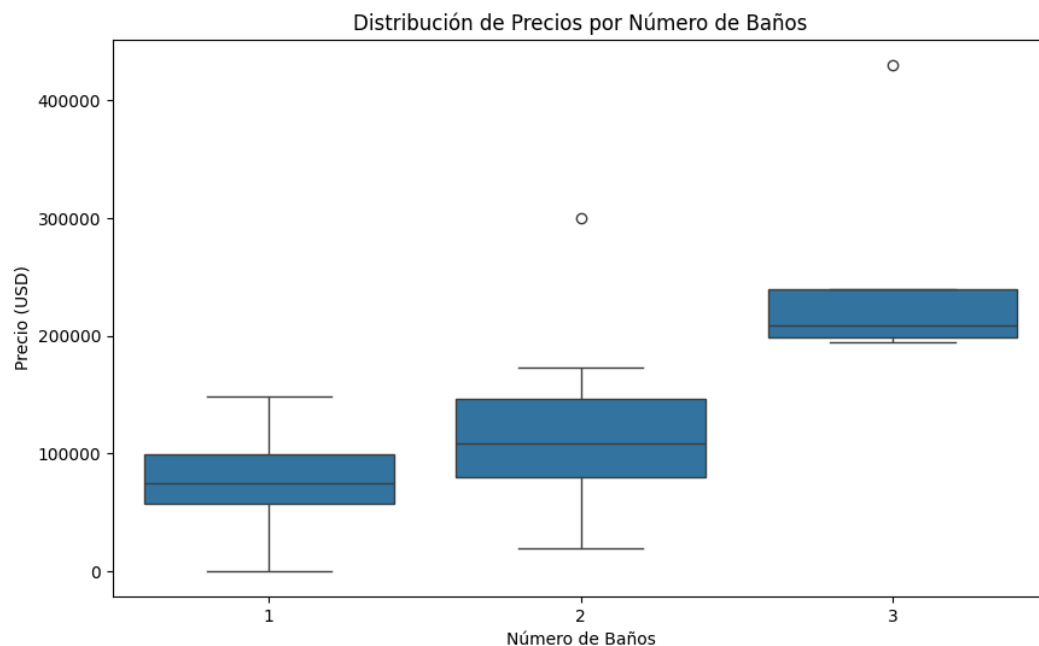
Conclusiones:

- **Mayor Precio con Más Ambientes:** Las propiedades con 6 o más ambientes tienen precios más altos, aunque la dispersión de precios dentro de este grupo sugiere que algunas propiedades pueden estar en ubicaciones menos exclusivas.
- **Opciones Accesibles en Propiedades con Menos Ambientes:** Las propiedades con 3 o 4 ambientes tienen precios más asequibles, lo que las hace atractivas para compradores que buscan un balance entre espacio y costo.

Recomendación: Promocionar propiedades con 6 o más ambientes como opciones de lujo para familias grandes, mientras que las propiedades con 3 o 4 ambientes pueden enfocarse en parejas jóvenes o pequeñas familias.

6. Distribución de Precios por Número de Baños

```
plt.figure(figsize=(10, 6))
sns.boxplot(x='Baños', y='Precio', data=df_remax)
plt.title('Distribución de Precios por Número de Baños')
plt.xlabel('Número de Baños')
plt.ylabel('Precio (USD)')
plt.savefig('images/Remax/Remax_exp_precio_por_banos.png')
plt.show()
```



Descripción: Este gráfico de caja ilustra cómo varían los precios de las propiedades en función del número de baños que tienen. Los precios de propiedades con uno, dos o tres baños se comparan para identificar patrones en la fijación de precios.

Observación General: A medida que aumenta el número de baños, los precios también tienden a incrementarse. Las propiedades con tres baños o más tienen precios significativamente más altos, lo que sugiere que los compradores perciben estas propiedades como más valiosas o lujosas.

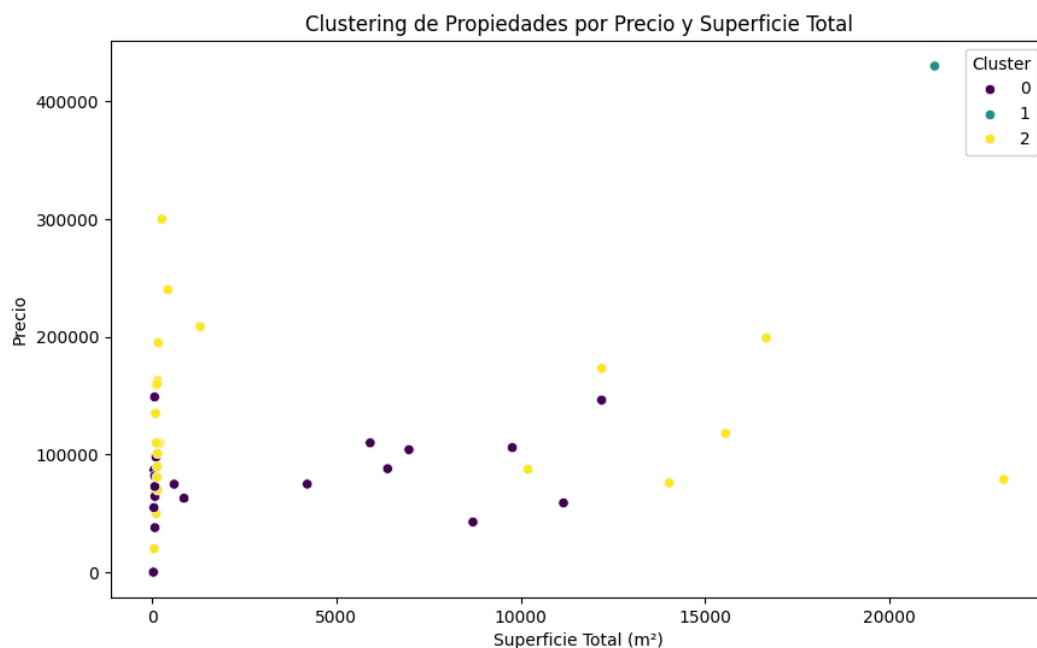
Conclusiones:

- **Mayor Valor con Más Baños:** Las propiedades con tres baños son las más costosas, con algunos valores atípicos que llegan a precios superiores a los \$300,000.
- **Segmento Medio con Dos Baños:** Las propiedades con dos baños parecen estar en un punto intermedio, atrayendo a compradores que buscan un equilibrio entre precio y comodidad.

Recomendación: Enfocar la comercialización de propiedades con dos baños hacia familias que buscan un balance entre precio y funcionalidad, mientras que las propiedades con tres baños pueden ser ofrecidas como opciones premium o de lujo.

7. Clustering de Propiedades según Características Clave

```
# Seleccionar las características relevantes
features = df_remax[['Precio', 'Superficie Total (m²)', 'Baños',
'Ambientes']]
# Escalar los datos
scaler = StandardScaler()
scaled_features = scaler.fit_transform(features)
# Aplicar K-Means
kmeans = KMeans(n_clusters=3, random_state=42)
df_remax['Cluster'] = kmeans.fit_predict(scaled_features)
# Visualizar los clusters
plt.figure(figsize=(10, 6))
sns.scatterplot(x='Superficie Total (m²)', y='Precio',
hue='Cluster', data=df_remax, palette='viridis')
plt.title('Clustering de Propiedades por Precio y Superficie Total')
plt.savefig('images/Remax/Remax_exp_clustering.png')
plt.show()
```



Descripción: Este gráfico muestra el agrupamiento de las propiedades en función de su precio y superficie total. Cada punto representa una propiedad, y los colores indican a qué cluster pertenece cada una.

Observación General: Se identifican tres clusters distintos de propiedades, que reflejan diferentes rangos de precio y superficie. El primer cluster agrupa propiedades de precios más bajos y superficies reducidas, mientras que el tercer cluster incluye propiedades más grandes y costosas.

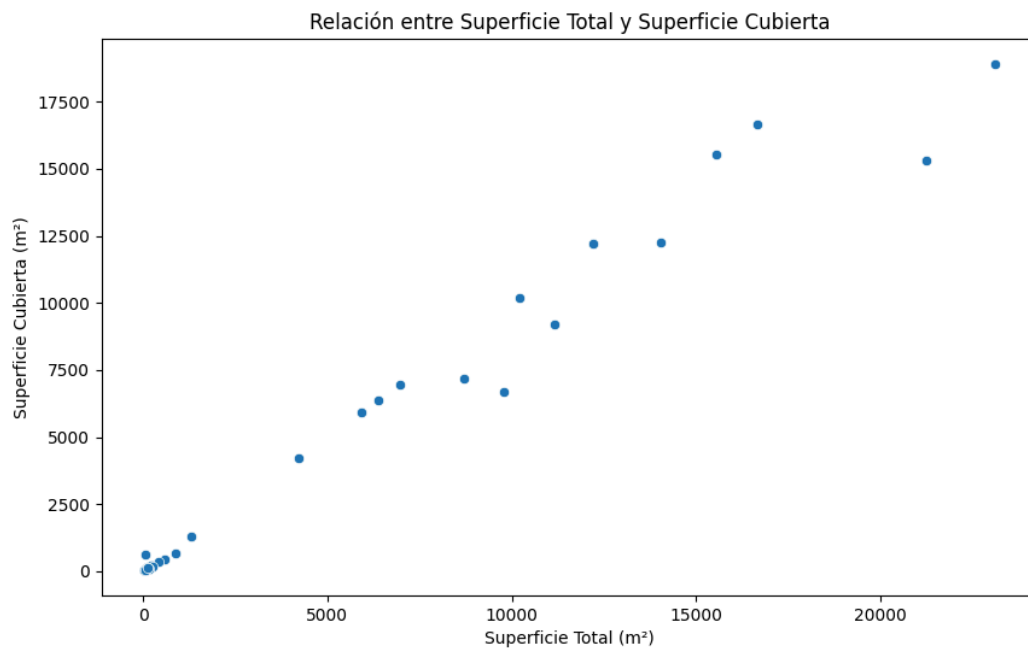
Conclusiones:

- **Segmentación del Mercado:** El mercado de Remax parece estar dividido en tres segmentos principales: propiedades accesibles, propiedades de rango medio y propiedades de lujo. Los compradores que buscan propiedades pequeñas tienen opciones dentro de un rango de precio más bajo, mientras que las propiedades más grandes y exclusivas pertenecen a un nicho de mercado de lujo.
- **Oportunidad en el Segmento Medio:** El segundo cluster, que incluye propiedades con superficies intermedias y precios en el rango medio, es un segmento interesante para explotar, ya que parece atraer a una base de compradores más amplia.

Recomendación: Crear estrategias de marketing diferenciadas para cada uno de estos segmentos. Las propiedades en el primer cluster deben promocionarse para compradores primerizos, mientras que las del tercer cluster deben enfocarse hacia un público con mayor capacidad adquisitiva. El segmento intermedio es clave para obtener un volumen alto de ventas.

8. Análisis de Discrepancia entre Superficie Total y Cubierta

```
plt.figure(figsize=(10, 6))
sns.scatterplot(x='Superficie Total (m²)', y='Superficie Cubierta (m²)', data=df_remax)
plt.title('Relación entre Superficie Total y Superficie Cubierta')
plt.xlabel('Superficie Total (m²)')
plt.ylabel('Superficie Cubierta (m²)')
plt.savefig('images/Remax/Remax_exp_superficie_total_vs_cubierta.png')
plt.show()
```



Descripción: Este gráfico de dispersión muestra la correlación entre la superficie total de una propiedad (que incluye espacios abiertos o sin construir) y la superficie cubierta (el área construida o habitable).

Observación General: Existe una fuerte correlación positiva entre la superficie total y la superficie cubierta. Sin embargo, se observan algunas propiedades con grandes áreas totales, pero superficies cubiertas significativamente más pequeñas, lo que sugiere que parte de la superficie es espacio no construido (jardines, patios, etc.).

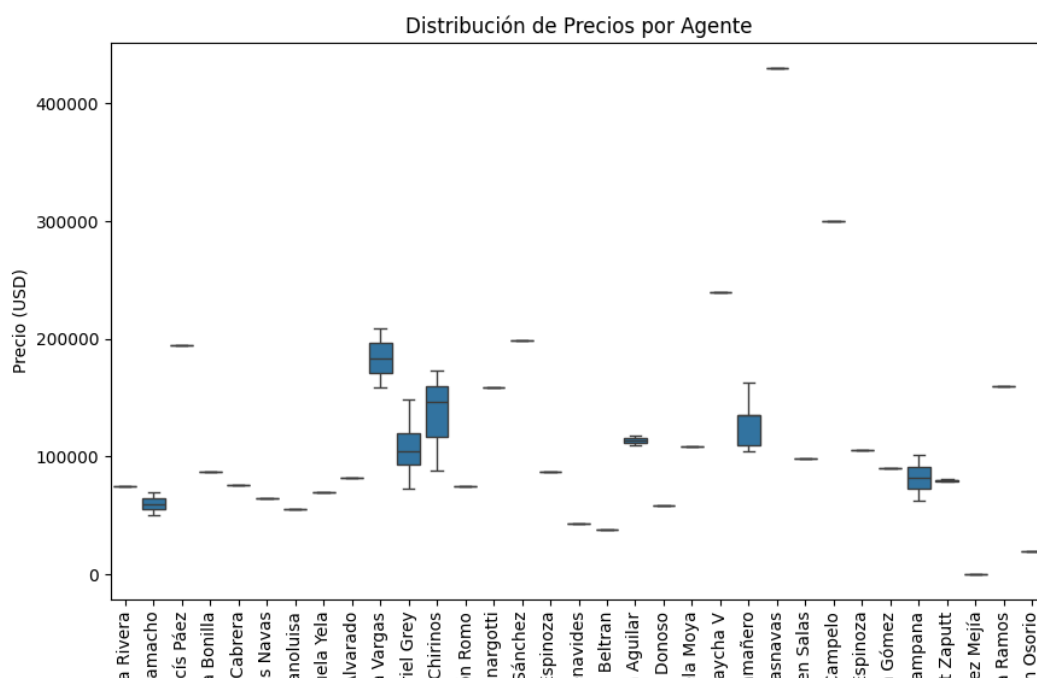
Conclusiones:

- **Propiedades Grandes:** Las propiedades más grandes, en términos de superficie total, tienden a tener también superficies construidas amplias, lo que puede estar asociado a un mayor precio.
- **Espacios Abiertos:** Propiedades con grandes superficies no cubiertas pueden incluir terrenos con jardines, áreas de recreación o espacios no edificadas, lo que podría ser atractivo para ciertos compradores que valoren espacios exteriores.

Recomendación: Las propiedades con grandes superficies abiertas deben destacarse en campañas de marketing para compradores que buscan viviendas con jardines o terrenos adicionales. Este tipo de propiedades podría atraer a familias o compradores interesados en el potencial de expansión.

9. Distribución de Precios por Agente

```
plt.figure(figsize=(10, 6))
sns.boxplot(x='Agente', y='Precio', data=df_remax)
plt.xticks(rotation=90)
plt.title('Distribución de Precios por Agente')
plt.xlabel('Agente')
plt.ylabel('Precio (USD)')
plt.savefig('images/Remax/Remax_exp_precio_por_agente.png')
plt.show()
```



Descripción: El gráfico de caja muestra cómo varían los precios de las propiedades listadas por cada agente. Los valores en el eje y representan el precio total de las propiedades, mientras que los diferentes agentes están listados en el eje x.

Observación General: Existe una disparidad significativa entre los precios manejados por los distintos agentes. Algunos agentes se especializan en propiedades de bajo costo, mientras que otros trabajan con propiedades más exclusivas y de mayor precio.

Conclusiones:

- **Variedad de Precios:** Algunos agentes parecen estar especializados en propiedades de lujo, con precios que llegan a superar los \$400,000, mientras que otros manejan mayormente propiedades más asequibles, en el rango de \$50,000 a \$150,000.
- **Dispersión de Precios:** La dispersión de precios dentro de los agentes sugiere que, aunque algunos tienen un enfoque claro en ciertos segmentos del mercado, otros

manejan propiedades en rangos de precios más variados, adaptándose a las necesidades de diferentes tipos de compradores.

Recomendación: Aprovechar esta información para crear perfiles de agentes especializados en ciertos segmentos del mercado, permitiendo estrategias de venta más personalizadas. Además, podría ser útil asignar agentes a propiedades en función de su especialización para maximizar la eficiencia de las ventas.

10. Análisis de Palabras Clave en Descripciones

```
# Lista de palabras a eliminar
palabras_a_eliminar = ['departamento', 'venta', 'se vende', 'venta',
'de', 'en', 'con', 'el']
regex_patron = re.compile(r'\b(?:' + '|'.join(palabras_a_eliminar) +
r')\b', flags=re.IGNORECASE)
# Unir todas las descripciones y eliminar las palabras no deseadas
texto_completo = ' '.join(df_remax['Descripción'].dropna())
texto_filtrado = regex_patron.sub('', texto_completo)
# Generar la nube de palabras
wordcloud = WordCloud(width=800, height=400,
background_color='white').generate(texto_filtrado)
# Mostrar la nube de palabras
plt.figure(figsize=(10, 5))
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis('off')
plt.title('Palabras Más Comunes en Descripciones de Propiedades')
plt.savefig('images/Remax/Remax_exp_wordcloud_descripciones.png')
plt.show()
```



Descripción: Este gráfico refleja las palabras más recurrentes en las descripciones de propiedades. El tamaño de las palabras indica la frecuencia con la que aparecen en las descripciones, destacando los conceptos más mencionados.

Observación General: Las palabras más frecuentes en las descripciones incluyen “Suite”, “Sector”, “Dormitorios” y “Conjunto”. Esto sugiere que el mercado está muy orientado a la venta de suites y apartamentos en sectores específicos de la ciudad. Además, la repetición de palabras como "estrenar" indica una presencia considerable de propiedades nuevas o renovadas.

Conclusiones:

- **Frecuencia de Suites:** El término "Suite" destaca considerablemente, lo que indica una oferta sustancial de este tipo de propiedades, posiblemente dirigidas a solteros, parejas jóvenes o profesionales que buscan propiedades más compactas.
- **Sector y Dormitorios:** Las descripciones también mencionan frecuentemente la ubicación (sector) y el número de dormitorios, lo cual sugiere que estas son características clave para los compradores en el mercado de Remax.
- **Nuevas Propiedades:** Palabras como "estrenar" o "amoblado" reflejan un mercado donde los compradores están buscando tanto propiedades nuevas como propiedades listas para habitar.

Recomendación: Considerar campañas de marketing enfocadas en suites y apartamentos con múltiples dormitorios. Además, podría ser beneficioso resaltar propiedades nuevas o recién renovadas, dado el interés mostrado por este tipo de características.

Conclusión General del Análisis Exploratorio de Remax:

El análisis exploratorio de las propiedades listadas en el portal Remax revela varias tendencias importantes en el mercado inmobiliario de Quito. Este mercado presenta características específicas que pueden aprovecharse para optimizar estrategias de venta y marketing, así como para identificar oportunidades de inversión.

1. Relación entre Precio y Superficie Total:

Como es esperable, las propiedades más grandes tienden a tener precios más altos. Sin embargo, existen casos en los que propiedades con superficies grandes se ofrecen a precios más bajos, lo que sugiere que el precio no solo depende del tamaño. Este fenómeno puede estar vinculado a factores como la ubicación, antigüedad de la propiedad o el estado de la infraestructura. Por otro lado, se identifican algunas propiedades con superficies pequeñas pero precios altos, lo que claramente indica que la ubicación es un factor clave en el valor de mercado, siendo las zonas más demandadas aquellas que elevan el costo de propiedades más pequeñas.

2. Distribución de Precios por Agente:

El análisis de la distribución de precios por agente revela que algunos agentes manejan un portafolio con precios más estandarizados, mientras que otros tienen una gama más amplia de propiedades, tanto en términos de precio como de características. Los agentes con una mayor dispersión de precios suelen estar relacionados con propiedades en diferentes segmentos de mercado, desde aquellas dirigidas a la clase media hasta propiedades de lujo. Esto muestra que los agentes con mayor diversificación pueden acceder a una mayor audiencia, pero también deben tener habilidades para gestionar propiedades de diferentes gamas.

3. Distribución de Precios por Número de Baños y Ambientes:

Las propiedades con más de 2 baños tienden a tener precios más elevados, lo que sugiere que el número de baños es un indicador significativo del nivel de lujo o confort de una propiedad. En cuanto a los ambientes, se observa que las propiedades con más ambientes suelen tener precios más elevados, aunque en este caso la variabilidad de precios es menor. El número de baños parece ser un factor más importante para los compradores a la hora de determinar el valor percibido de una propiedad, en comparación con el número de ambientes.

4. Distribución de Precios y Gastos:

Existe una correlación positiva entre las expensas mensuales y el precio de la propiedad. Esto es lógico, ya que las propiedades más costosas suelen estar ubicadas en edificios o urbanizaciones con mayores servicios, lo que se refleja en expensas más altas. No obstante, también se detectan algunas propiedades con expensas relativamente bajas en comparación con su precio, lo que podría ser un punto de venta atractivo para compradores que buscan minimizar los costos de mantenimiento.

5. Palabras Clave en las Descripciones de Propiedades:

El análisis de palabras clave en las descripciones de las propiedades muestra que términos como "Suite", "Dormitorios", "Sector" y "Conjunto" son recurrentes, lo que refleja una preferencia por describir propiedades en función de su ubicación y sus características físicas. Términos como "Estrenar" o "Amoblado" también son comunes, lo que sugiere que hay una demanda considerable por propiedades

nuevas o que ya están listas para habitar. Estas descripciones tienden a resaltar las características diferenciadoras para atraer a los compradores, como la ubicación estratégica o el estado de la propiedad.

6. **Clustering de Propiedades:**

El análisis de clustering sugiere que existen tres grandes segmentos de mercado en términos de precio y tamaño de las propiedades. Un grupo está compuesto por propiedades pequeñas y económicas, dirigidas principalmente a compradores primerizos o a quienes buscan una opción accesible. Otro grupo está compuesto por propiedades medianas, con precios moderados, orientadas a familias de clase media. El tercer grupo, más reducido, incluye propiedades de lujo, que tienen precios significativamente más altos y, por lo general, mayor superficie. Esta segmentación es clave para personalizar estrategias de marketing y ventas para cada uno de estos grupos.

7. **Distribución de Superficies y Correlación con Precios:**

La mayoría de las propiedades en el dataset tienen una superficie total menor a 5,000 m², lo que sugiere que el mercado está dominado por propiedades de tamaño mediano. Las propiedades más grandes, por encima de los 10,000 m², son más escasas y tienden a estar en el segmento de lujo. La correlación entre el precio y la superficie total es positiva, como es lógico, pero la dispersión observada en el análisis sugiere que la ubicación y las características adicionales, como las amenidades, también juegan un papel crucial en el precio final.

8. **Distribución de Precios:**

El mercado está dominado por propiedades en el rango de \$50,000 a \$150,000, lo que refleja un enfoque en compradores de clase media. Las propiedades de lujo, con precios superiores a \$300,000, son menos comunes, pero existen. Esta segmentación del mercado indica que las estrategias de marketing deben centrarse principalmente en atraer a compradores de clase media, pero con un enfoque en destacar las propiedades de lujo como oportunidades únicas para aquellos con mayor poder adquisitivo.

9. **Correlaciones entre Variables:**

El análisis de correlaciones entre variables muestra que el precio de una propiedad está moderadamente relacionado con la superficie total y con el número de baños. Aunque estas relaciones no son extremadamente fuertes, indican que los compradores tienden a valorar las propiedades más grandes y con más comodidades, pero otros factores, como la ubicación y las amenidades, también juegan un papel importante. Esto sugiere que una combinación de varias características es clave para definir el valor de mercado de una propiedad.