

Diccionario

Nombre Columna		Descripción
Comuna	Comuna	Comuna en la que se encuentra la propiedad
Propiedad	Propiedad	Tipo de propiedad (Casa o Departamento)
Precio	Precio	Precio de la propiedad
Direccion	Dirección	Dirección "mas" exacta de la propiedad
Tamanios	Tamaño	Tamaño en m2 de la propiedad
Dormitorios	Número de dormitorios	Número de dormitorios que posee la propiedad
Baños	Número de baños	Número de baños que posee la propiedad

In [1]:

```
import matplotlib.pyplot as plt
import pandas as pd

df = pd.read_excel("propiedadesgranconcepcion.xlsx")
df
```

Out[1]:

	Comuna	Propiedad	Precio	Direccion	Tamanios	Dormitorios	Baños
0	['Tomé, Biobío']	Casa	5.990	Pingual, Tomé, Biobío	['559 m² totales']	['4 dormitorios']	['2 baños']
1	['Tomé, Biobío']	Casa	7.900	Avenida Pingual, Tomé, Chile, Pingual, Tom...	['1068 m² totales']	['4 dormitorios']	['4 baños']
2	['Tomé, Biobío']	Casa	8.120	Bella Casa Con Piscina 3d 2b T350 M2 , Tomé, T...	['350 m² totales']	['3 dormitorios']	['2 baños']
3	['Tomé, Biobío']	Casa	6.900	Avenida Pingual, Tomé, Chile, Pingual, Tom...	['560 m² totales']	['4 dormitorios']	['3 baños']
4	['Tomé, Biobío']	Casa	33.700	Avenida Pingual, Tomé, Chile, Pingual, Tom...	['875 m² totales']	['6 dormitorios']	['4 baños']
...	...	...	...	...	...	...	...
3110	['Chiguayante, Biobío']	Departamento	2.373	Esperanza 507, Chiguayante, Chile, La Leonera,...	[]	[]	[]
3111	['Chiguayante, Biobío']	Departamento	5.000	La Leonera, Chiguayante, Biobío	['85 m² totales']	['2 dormitorios']	['2 baños']
3112	['Chiguayante, Biobío']	Departamento	74.000.000	Esperanza 507, Chiguayante, Chile, La Leonera,...	['59 m² totales']	['3 dormitorios']	['1 baño']
3113	['Chiguayante, Biobío']	Departamento	350.000.000	Sanhueza 55, Concepción, Chile, Lonco, Chiguay...	['145 m² totales']	['4 dormitorios']	['3 baños']
3114	['Chiguayante, Biobío']	Departamento	4.100	Fresia 90, Chiguayante, Chile, Altos de Chigua...	['90 m² totales']	['3 dormitorios']	['2 baños']

3115 rows x 7 columns

Preprocesamiento

In [2]:

```
#Verificamos el tipo de dato de las variables, la cantidad de datos y si existen valores nulos

df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3115 entries, 0 to 3114
```

rangeindex: 3115 entries, 0 to 3114  
Data columns (total 7 columns):  
# Column Non-Null Count Dtype  
--- --  
0 Comuna 3115 non-null object  
1 Propiedad 3115 non-null object  
2 Precio 3115 non-null object  
3 Direccion 3115 non-null object  
4 Tamanios 3115 non-null object  
5 Dormitorios 3115 non-null object  
6 Baños 3115 non-null object  
dtypes: object(7)  
memory usage: 170.5+ KB

In [3]:

```
#Observamos que la variable Tamanios se refiere a Tamaño por lo que se modificará su nombre  
df.rename(columns={'Tamanios': 'Tamaño en m2'}, inplace=True)
```

In [4]:

```
#Observamos si existen filas duplicadas  
df.duplicated().sum()
```

Out[4]:  
38

In [5]:

```
#Borramos las filas duplicadas  
df.drop_duplicates()
```

Out[5]:

	Comuna	Propiedad	Precio	Direccion	Tamaño en m2	Dormitorios	Baños
0	['Tomé, Biobío']	Casa	5.990	Pingual, Tomé, Biobío	['559 m2 totales']	['4 dormitorios']	['2 baños']
1	['Tomé, Biobío']	Casa	7.900	Avenida Pingual, Tomé, Chile, Pingual, Tom...	['1068 m2 totales']	['4 dormitorios']	['4 baños']
2	['Tomé, Biobío']	Casa	8.120	Bella Casa Con Piscina 3d 2b T350 M2 , Tomé, T...	['350 m2 totales']	['3 dormitorios']	['2 baños']
3	['Tomé, Biobío']	Casa	6.900	Avenida Pingual, Tomé, Chile, Pingual, Tom...	['560 m2 totales']	['4 dormitorios']	['3 baños']
4	['Tomé, Biobío']	Casa	33.700	Avenida Pingual, Tomé, Chile, Pingual, Tom...	['875 m2 totales']	['6 dormitorios']	['4 baños']
...	...	...	...	...	...	...	...
3110	['Chiguayante, Biobío']	Departamento	2.373	Esperanza 507, Chiguayante, Chile, La Leonera,...	[]	[]	[]
3111	['Chiguayante, Biobío']	Departamento	5.000	La Leonera, Chiguayante, Biobío	['85 m2 totales']	['2 dormitorios']	['2 baños']
3112	['Chiguayante, Biobío']	Departamento	74.000.000	Esperanza 507, Chiguayante, Chile, La Leonera,...	['59 m2 totales']	['3 dormitorios']	['1 baño']
3113	['Chiguayante, Biobío']	Departamento	350.000.000	Sanhueza 55, Concepción, Chile, Lonco, Chiguy...	['145 m2 totales']	['4 dormitorios']	['3 baños']
3114	['Chiguayante, Biobío']	Departamento	4.100	Fresia 90, Chiguayante, Chile, Altos de Chigua...	['90 m2 totales']	['3 dormitorios']	['2 baños']

3077 rows x 7 columns

In [6]:

```
In [6]:
```

```
#Buscamos si en la variable Dormitorios se encuentra la letra "a"
```

```
filas_con_a = df[df['Dormitorios'].str.contains('a', case=False)]
```

```
print(filas_con_a)
```

	Comuna	Propiedad	Precio \
97	['Tomé, Biobío']	Departamento	5.490
111	['Talcahuano, Biobío']	Casa	5.550
204	['Talcahuano, Biobío']	Casa	4.400
215	['Talcahuano, Biobío']	Casa	160.000.000
221	['Talcahuano, Biobío']	Casa	3.500
...	...	...	...
2749	['Concepción, Biobío']	Departamento	1.590
2760	['Concepción, Biobío']	Departamento	2.241
2893	['Chiguayante, Biobío']	Casa	3.500
3018	['Chiguayante, Biobío']	Departamento	2.900
3019	['Chiguayante, Biobío']	Departamento	2.911

	Direccion \
97	Pedro Leon Ugalde 45, Dichato, Tomé, Chile, Di...
111	Calle Tierras Coloradas S/n, Talcahuano, Biobío
204	Calle Nueva 100, Talcahuano, Chile, Talcahuano...
215	Condominio Cipreces De Malaga, Carriel, Talcah...
221	Claudio Gay 100, Talcahuano, Chile, Carriel, T...
...	...
2749	Colo Colo, Concepción, Chile, Barrio Poniente,...
2760	Calle Maipú 1740, Barrio Oriente - Las Tres Pa...
2893	Pinares, Chiguayante Sur, Chiguayante, Biobío
3018	Obispo San Miguel 1841, Chiguayante, Chile, Al...
3019	Cornelio Saavedra 11, Altos de Chiguayante - P...

	Tamaño en m2	Dormitorios	Baños
97	['2 dormitorios']	['2 baños']	['']
111	['85.8 a 128.42 m² útiles']	['3 a 4 dormitorios']	['3 baños']
204	['3 dormitorios']	['2 baños']	['']
215	['4 dormitorios']	['3 baños']	['']
221	['4 dormitorios']	['2 baños']	['']
...	...	...	...
2749	['33 m² totales']	['1 baño']	['']
2760	['22.08 a 55.31 m² útiles']	['1 a 2 dormitorios']	['1 a 2 baños']
2893	['3 dormitorios']	['3 baños']	['']
3018	['47.22 a 73.7 m² útiles']	['2 a 3 dormitorios']	['2 baños']
3019	['35.61 a 92.19 m² útiles']	['1 a 3 dormitorios']	['1 a 2 baños']

```
[104 rows x 7 columns]
```

Como podemos observar al buscar las filas de la variable Dormitorios que contengan la letra "a" obtenemos dos tipos de resultados siendo estos por ejemplo: ['3 a 4 dormitorios'] o ['2 baños'], en el primer caso esto ocurre por que al hacer webscraping habían anuncios de proyectos inmobiliarios donde existían distintos modelos de propiedades y en el segundo caso ocurre cuando faltan algunos datos y la información de la variable Baño se corre un lugar hacia al lado. Estos dos casos se eliminarán ya que de los proyectos inmobiliarios no hay mayor información y por otro lado se necesita tener la información completa.

```
In [7]:
```

```
#Dejamos en nuestro data frame solo las filas que NO contienen la letra "a" en la variable Dormitorio
```

```
df = df[~df['Dormitorios'].str.contains('a')]
```

Como se observa en el inicio todas las variables son de tipo "object", sin embargo las variables Tamaño en m2, Dormitorios, Baños, Precio, son valores numéricos por lo que se deben cambiar a "int" pero antes de esto se deben eliminar los caracteres y palabras que se encuentran en estas variables.

```
In [8]:
```

```
#En la variable Comuna se deja solo el nombre de esta eliminando caracteres y el nombre d
```

*e la región en las demas variables se eliminan los caracteres dejando solo el número*

```
df['Comuna'] = df['Comuna'].str.split(',').str[0]
df['Comuna'] = df['Comuna'].str.split('\\').str[-1]
df['Tamaño en m2'] = df['Tamaño en m2'].str.extract('(\\d+)')
df['Dormitorios'] = df['Dormitorios'].str.extract('(\\d+)')
df['Baños'] = df['Baños'].str.extract('(\\d+)')
```

```
C:\Users\danie\AppData\Local\Temp\ipykernel_36532\2098561331.py:3: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)

```
df['Comuna'] = df['Comuna'].str.split(',').str[0]
```

```
C:\Users\danie\AppData\Local\Temp\ipykernel_36532\2098561331.py:4: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)

```
df['Comuna'] = df['Comuna'].str.split('\\').str[-1]
```

```
C:\Users\danie\AppData\Local\Temp\ipykernel_36532\2098561331.py:5: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)

```
df['Tamaño en m2'] = df['Tamaño en m2'].str.extract('(\\d+)')
```

```
C:\Users\danie\AppData\Local\Temp\ipykernel_36532\2098561331.py:6: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)

```
df['Dormitorios'] = df['Dormitorios'].str.extract('(\\d+)')
```

```
C:\Users\danie\AppData\Local\Temp\ipykernel_36532\2098561331.py:7: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)

```
df['Baños'] = df['Baños'].str.extract('(\\d+)')
```

In [9]:

```
#Luego de estas modificaciones observamos si existen valores nulos.
```

```
df.isna().sum().sum()
```

Out[9]:

55

In [10]:

```
#Se eliminan filas que poseen valores nulos.
```

```
df = df.dropna()
```

In [11]:

```
#Se modifica el tipo de variables de Precio, Dormitorios, Tamaño en m2 y Baños pasando a "int".
#Antes de modificar la variable Precio se debió eliminar los puntos que poseían los numeros.
```

```
df['Precio'] = df['Precio'].str.replace('.', '', regex=True).astype(int)
df['Precio'] = pd.to_numeric(df['Precio'], errors='coerce')
df['Dormitorios'] = pd.to_numeric(df['Dormitorios'], errors='coerce')
df['Tamaño en m2'] = pd.to_numeric(df['Tamaño en m2'], errors='coerce')
df['Baños'] = pd.to_numeric(df['Baños'], errors='coerce')
```

In [12]:

```
df.info()
df.head()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 2981 entries, 0 to 3114
Data columns (total 7 columns):
#   Column          Non-Null Count  Dtype
---  -
0   Comuna          2981 non-null   object
1   Propiedad       2981 non-null   object
2   Precio          2981 non-null   int32
3   Direccion       2981 non-null   object
4   Tamaño en m2    2981 non-null   int64
5   Dormitorios     2981 non-null   int64
6   Baños           2981 non-null   int64
dtypes: int32(1), int64(3), object(3)
memory usage: 174.7+ KB
```

Out[12]:

	Comuna	Propiedad	Precio	Direccion	Tamaño en m2	Dormitorios	Baños
0	Tomé	Casa	5990	Pingüeral, Tomé, Biobío	559	4	2
1	Tomé	Casa	7900	Avenida Pingüeral, Tomé, Chile, Pingüeral, Tom...	1068	4	4
2	Tomé	Casa	8120	Bella Casa Con Piscina 3d 2b T350 M2 , Tomé, T...	350	3	2
3	Tomé	Casa	6900	Avenida Pingüeral, Tomé, Chile, Pingüeral, Tom...	560	4	3
4	Tomé	Casa	33700	Avenida Pingüeral, Tomé, Chile, Pingüeral, Tom...	875	6	4

En la Variable Precio hay valores que se encuentran en UF y otros en Pesos Chilenos (CLP) para una mejor comprension de estos se modificaran para dejarlos todos en Pesos Chilenos.

In [13]:

```
precios_menores_50000 = df.loc[df['Precio'] < 50000, 'Precio']
```

```
# Mostrar los valores resultantes
print(precios_menores_50000)
```

```
0      5990
1      7900
2      8120
3      6900
4     33700
...
3105    4900
3106    4500
3108    2650
3111    5000
3114    4100
Name: Precio, Length: 2143, dtype: int32
```

In [14]:

```
df.loc[df['Precio'] < 50000, 'Precio'] *= 36936.38
```

```
# Convertir la columna 'Precio' a enteros
df['Precio'] = df['Precio'].astype(int)

# Mostrar el DataFrame actualizado
```

```
print(df)
```

```
      Comuna      Propiedad      Precio \
0      Tomé      Casa      221248916
1      Tomé      Casa      291797402
2      Tomé      Casa      299923405
3      Tomé      Casa      254861021
4      Tomé      Casa      1244756006
...      ...      ...      ...
3109 Chiguayante Departamento      600000000
3111 Chiguayante Departamento      184681900
3112 Chiguayante Departamento      740000000
3113 Chiguayante Departamento      350000000
3114 Chiguayante Departamento      151439158

      Direccion      Tamaño en m2 \
0      Pingueral, Tomé, Biobío      559
1      Avenida Pingueral, Tomé, Chile, Pingueral, Tom...      1068
2      Bella Casa Con Piscina 3d 2b T350 M2 , Tomé, T...      350
3      Avenida Pingueral, Tomé, Chile, Pingueral, Tom...      560
4      Avenida Pingueral, Tomé, Chile, Pingueral, Tom...      875
...      ...      ...
3109 Esperanza 507, Chiguayante, Chile, La Leonera,...      53
3111      La Leonera, Chiguayante, Biobío      85
3112 Esperanza 507, Chiguayante, Chile, La Leonera,...      59
3113 Sanhueza 55, Concepción, Chile, Lonco, Chiguay...      145
3114 Fresia 90, Chiguayante, Chile, Altos de Chigua...      90

      Dormitorios      Baños
0      4      2
1      4      4
2      3      2
3      4      3
4      6      4
...      ...      ...
3109      3      1
3111      2      2
3112      3      1
3113      4      3
3114      3      2
```

```
[2981 rows x 7 columns]
```

**El último paso que se realizará en este preprocesamiento será buscar algunos valores anómalos, que en este caso radican del error humano de las personas al publicar los avisos de venta y como se mencionó mas arriba solo se buscan datos completos y "buenos"**

```
In [15]:
```

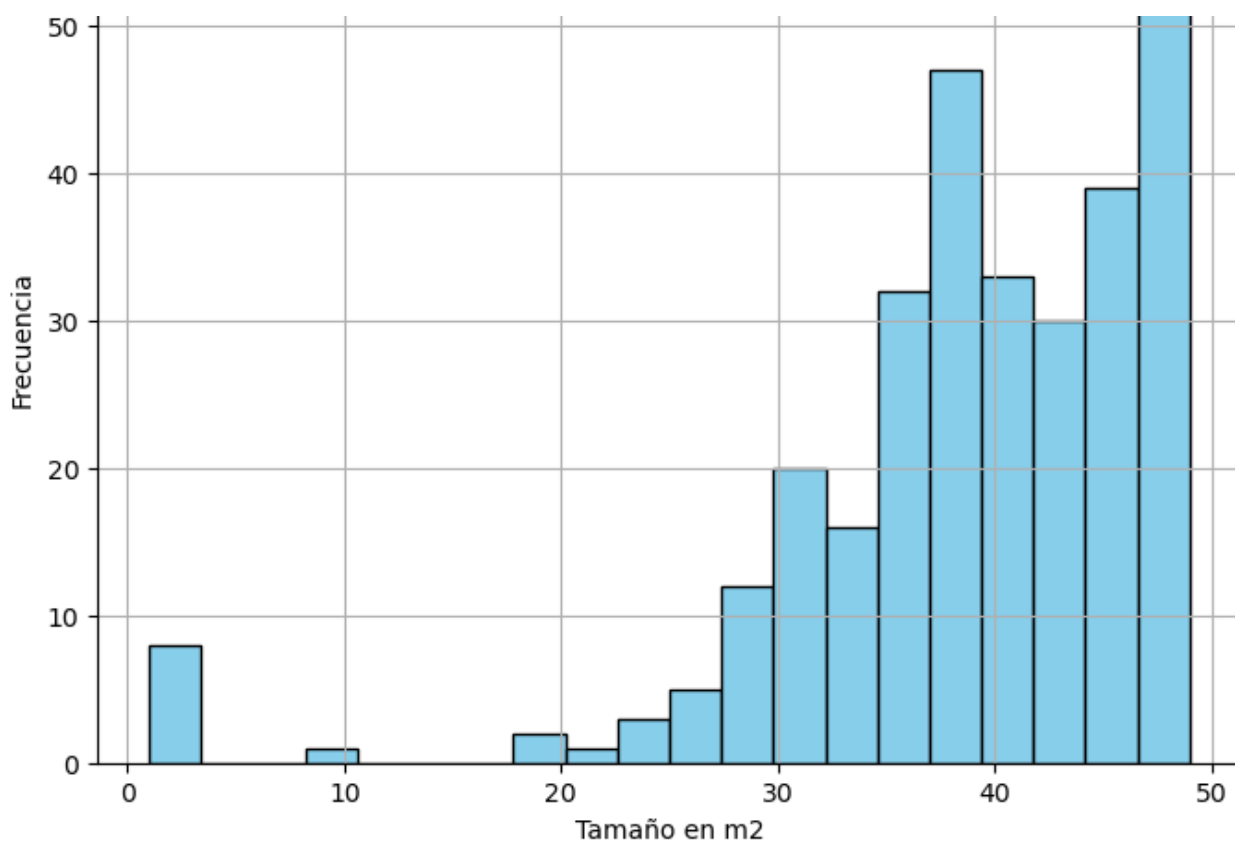
```
#Los primeros valores anómalos que se buscan son los de la variable Tamaño en m2, para es  
to se graficarán en primera instancia los valores menores a 50 m2  
#esto para poder ver con mayor detalle valores anómalos pequeños
```

```
df_filtrado = df[df['Tamaño en m2'] < 50]
```

```
plt.figure(figsize=(8, 6))
df_filtrado['Tamaño en m2'].plot.hist(color='skyblue', edgecolor='black', bins=20)
plt.title('Histograma de Tamaño en m2 (Valores menores a 50)')
plt.xlabel('Tamaño en m2')
plt.ylabel('Frecuencia')
plt.grid(True)
plt.show()
```

Histograma de Tamaño en m2 (Valores menores a 50)





Como se puede observar hay valores cercanos a 1 m2 por lo que se procederá a eliminar estos datos

In [16]:

```
#Se eliminan los valores en la variable "Tamaño en m2" menores a 15 m2 dejando así solo v  
alores mayores o iguales a 15 m2
```

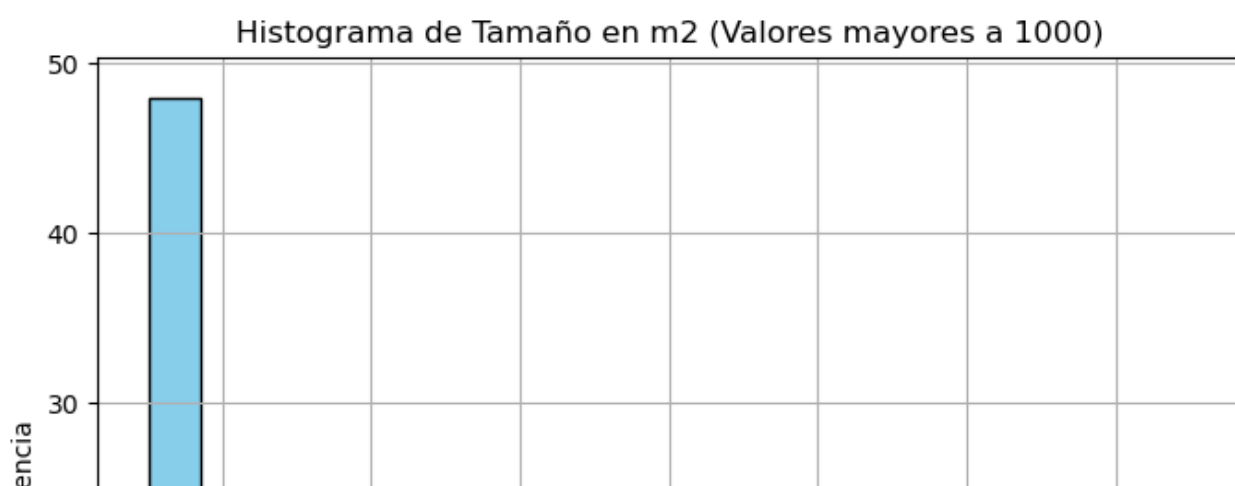
```
df = df[df['Tamaño en m2'] >= 15]
```

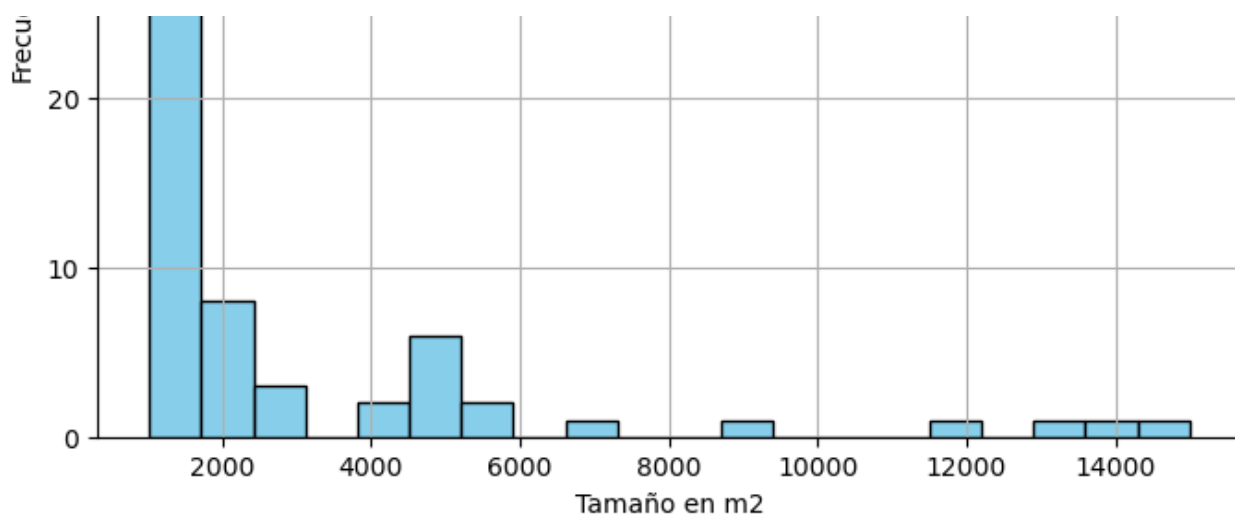
In [17]:

```
#Ahora se buscan valores anómalos mayores para esto se filtra los valores mayores a 1000  
para poder ver con mayor mayor detalle estos valores.
```

```
df_filtrado = df[df['Tamaño en m2'] > 1000]
```

```
plt.figure(figsize=(8, 6)) # Tamaño del gráfico  
df_filtrado['Tamaño en m2'].plot.hist(color='skyblue', edgecolor='black', bins=20) # Pl  
ot del histograma  
plt.title('Histograma de Tamaño en m2 (Valores mayores a 1000)') # Título del gráfico  
plt.xlabel('Tamaño en m2') # Etiqueta del eje x  
plt.ylabel('Frecuencia') # Etiqueta del eje y  
plt.grid(True) # Habilitar cuadrícula  
plt.show() # Mostrar el gráfico
```





En este caso no se consideran valores anómalos ya que pueden existir terrenos de esta cantidad de m2, se buscaban valores que se escaparan de lo normal como cientos de miles o millones.

```
In [18]:

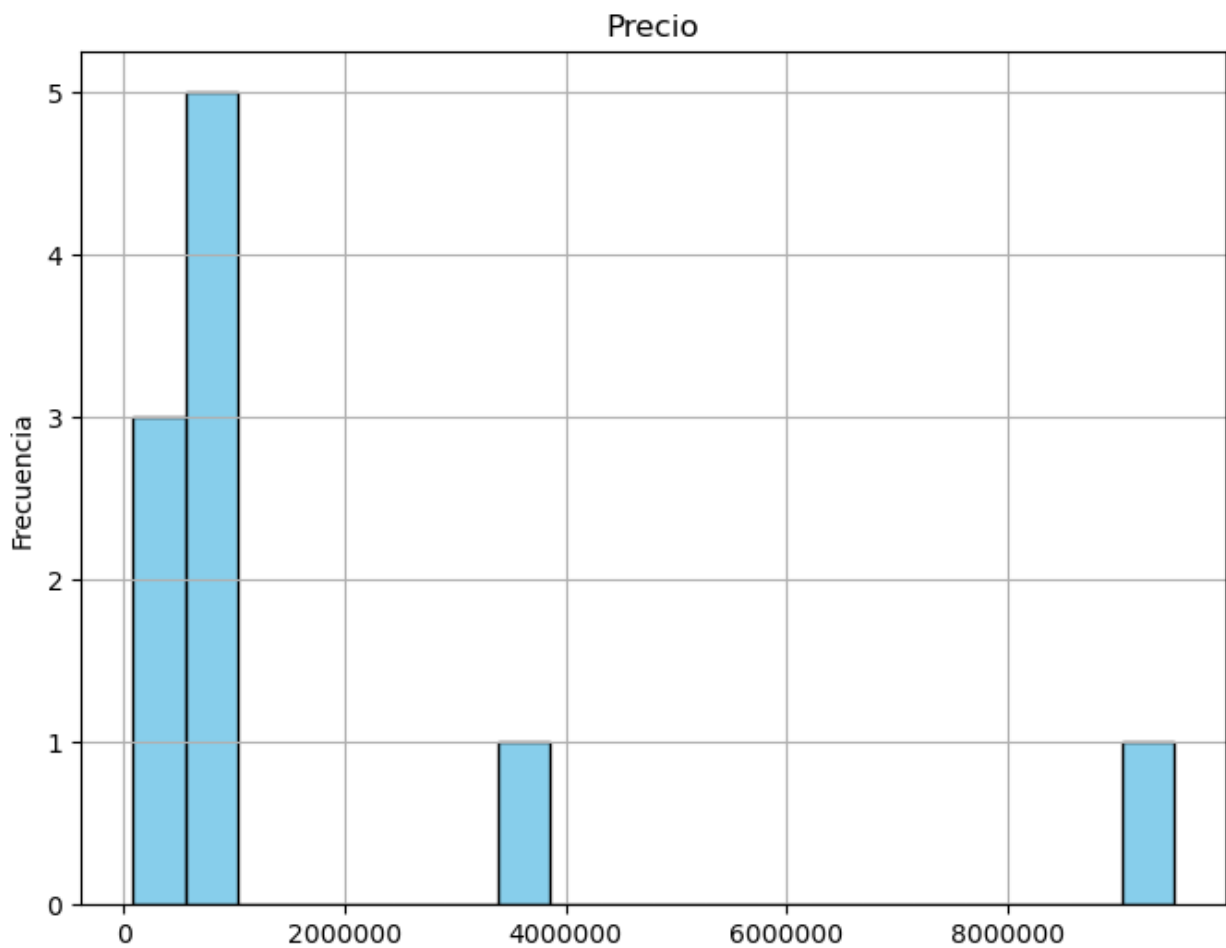
#Se realiza el mismo procedimiento con la variable "Precio", en este caso se filtran valores menores a $10.000.000

df_filtrado = df[df['Precio'] < 100000000]

plt.figure(figsize=(8, 6))
df_filtrado['Precio'].plot.hist(color='skyblue', edgecolor='black', bins=20)
plt.title('Precio')
plt.xlabel('Precio')
plt.ylabel('Frecuencia')
plt.grid(True)

# Cambiar el formato del eje x a notación regular
plt.ticklabel_format(style='plain', axis='x')

plt.show()
```





In [19]:

```
#En este caso también se buscan las filas que poseen estos valores pequeños para corroborar si existe error
```

```
errorprecio = df[df['Precio'] < 100000000]
```

```
# Mostrar las filas resultantes
print(errorprecio)
```

	Comuna	Propiedad	Precio \
129	Talcahuano	Casa	140000
301	Talcahuano	Departamento	850000
461	San Pedro de la Paz	Casa	82600
797	San Pedro de la Paz	Departamento	850000
904	San Pedro de la Paz	Departamento	200000
2015	Concepción	Departamento	9500000
2041	Concepción	Departamento	800000
2703	Concepción	Departamento	3500000
2761	Concepción	Departamento	701791
2762	Concepción	Departamento	600000

	Direccion	Tamaño en m2 \
129	Pichamávida, Sector Cruz Del Sur, Carriel, Tal...	300
301	Vasco Núñez De Balboa, Talcahuano, Chile, Carr...	71
461	Ernesto Pinto Lagarrigue, Km 1.5, San Pedro De...	14000
797	Avenida Los Parques, San Pedro De La Paz, Chil...	180
904	Avenida Costaner 7488 Edificio Bosque Mar, San...	140
2015	Angol 600, Concepción, Chile, Centro de Concep...	67
2041	Jorge Alessandri 560, 4090070 Concepción, Bío ...	55
2703	Freire 1165, Concepción, Chile, Centro de Conc...	38
2761	Lomas De San Sebastián, Concepción, Chile, Lom...	65
2762	Andalién, Concepción, Biobío	60

	Dormitorios	Baños
129	3	2
301	3	2
461	4	5
797	3	3
904	3	3
2015	3	2
2041	2	2
2703	1	1
2761	2	2
2762	3	2

**Viendo las características de las propiedades que tienen un valor menor a \$10.000.000 podemos decir que su valor se puede deber a un error de tipeo por lo que se borrarán estos datos.**

In [20]:

```
df = df[df['Precio'] >= 100000000]
```

In [21]:

```
#Ahora se grafican las variables Baños y Dormitorios
```

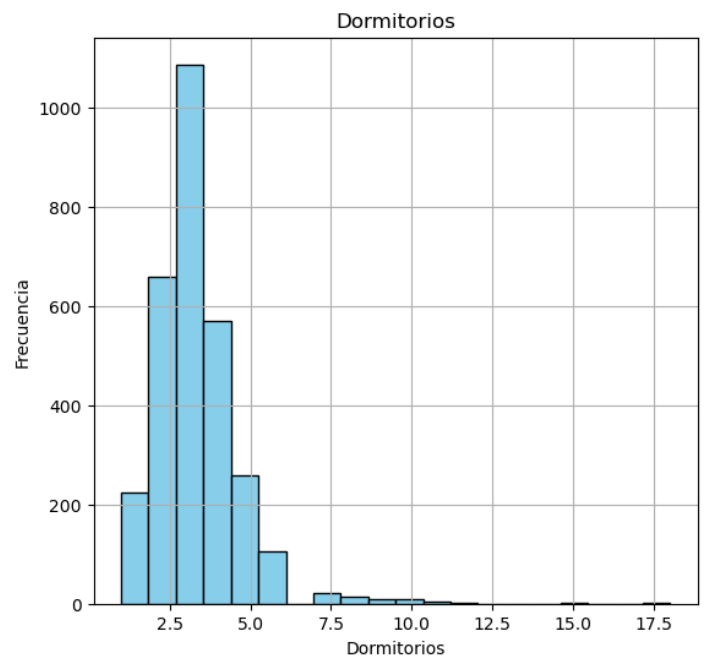
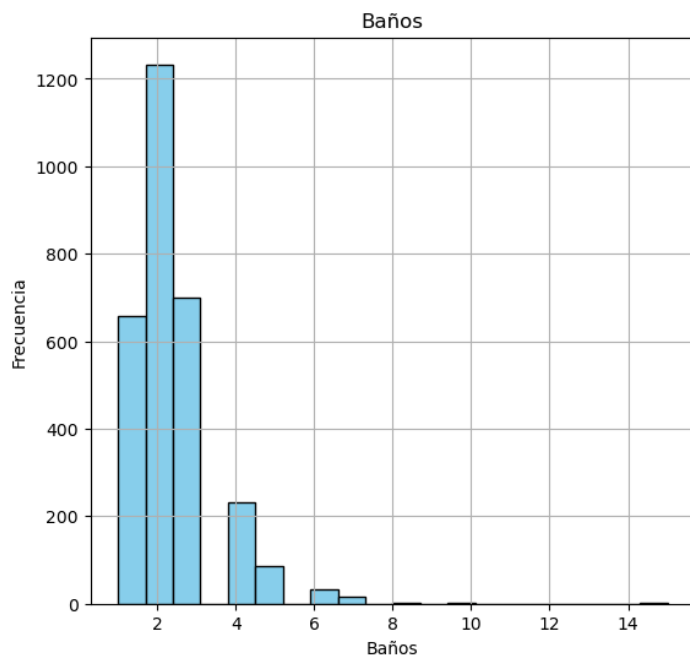
```
plt.figure(figsize=(14, 6))
```

```
#Baños
```

```
plt.subplot(1, 2, 1)
df['Baños'].plot.hist(color='skyblue', edgecolor='black', bins=20)
plt.title('Baños')
plt.xlabel('Baños')
plt.ylabel('Frecuencia')
plt.grid(True)
plt.ticklabel_format(style='plain', axis='x')
```

```
#Dormitorios
plt.subplot(1, 2, 2)
df['Dormitorios'].plot.hist(color='skyblue', edgecolor='black', bins=20)
plt.title('Dormitorios')
plt.xlabel('Dormitorios')
plt.ylabel('Frecuencia')
plt.grid(True)
plt.ticklabel_format(style='plain', axis='x')

plt.show()
```



Se pueden observar algunas propiedades que tienen una gran cantidad de baños y dormitorios en comparación a una vivienda "normal" por lo que se observarán todas las características de estos datos.

In [22]:

```
#Se filtran los datos de las propiedades que tienen mas de 6 dormitorios

errordormitorios = df[df['Dormitorios'] > 6]

# Mostrar las filas resultantes
print(errordormitorios)
```

	Comuna	Propiedad	Precio \
57	Tomé	Casa	697728218
60	Tomé	Casa	697728218
67	Tomé	Casa	110735267
124	Talcahuano	Casa	165000000
163	Talcahuano	Casa	130000000
166	Talcahuano	Casa	153800000
251	Talcahuano	Casa	339445332
256	Talcahuano	Casa	250000000
365	San Pedro de la Paz	Casa	690710306
372	San Pedro de la Paz	Casa	620531184
459	San Pedro de la Paz	Casa	690710306
522	San Pedro de la Paz	Casa	690710306
528	San Pedro de la Paz	Casa	583594804
559	San Pedro de la Paz	Casa	620531184
654	San Pedro de la Paz	Casa	690710306
710	San Pedro de la Paz	Casa	554045700
1212	Hualpé	Casa	89000000
1324	Coronel	Casa	315000000
1463	Concepción	Casa	616837546
1493	Concepción	Casa	701791220
1522	Concepción	Casa	250000000
1531	Concepción	Casa	757195790
1536	Concepción	Casa	454317473
1542	Concepción	Casa	720259410

1549	Concepción	Casa	400000000
1552	Concepción	Casa	420000000
1556	Concepción	Casa	313959230
1559	Concepción	Casa	428000000
1570	Concepción	Casa	572513890
1578	Concepción	Casa	340000000
1608	Concepción	Casa	489407034
1626	Concepción	Casa	517109319
1638	Concepción	Casa	1081127842
1640	Concepción	Casa	929775000
1652	Concepción	Casa	812600360
1673	Concepción	Casa	554045700
1685	Concepción	Casa	570496584
1694	Concepción	Casa	282858798
1707	Concepción	Casa	200000000
1722	Concepción	Casa	430000000
1725	Concepción	Casa	205000000
1762	Concepción	Casa	496683501
1784	Concepción	Casa	435849283
1791	Concepción	Casa	432155645
1798	Concepción	Casa	250000000
1801	Concepción	Casa	522649776
1803	Concepción	Casa	499269048
1811	Concepción	Casa	295491040
1828	Concepción	Casa	230000000
1829	Concepción	Casa	320000000
1840	Concepción	Casa	738727600
2408	Concepción	Departamento	443236559
2809	Chiguayante	Casa	1658443462
2822	Chiguayante	Casa	565126614
2852	Chiguayante	Casa	879085843
2853	Chiguayante	Casa	879085843
2862	Chiguayante	Casa	404453361
2897	Chiguayante	Casa	1658443462
2912	Chiguayante	Casa	590982080

	Dirección	Tamaño en m2	\
57	Sotomayor 1053, Tome, Tomé, Chile, Ciudad de T...	490	
60	Sotomayor 1053, Tome, Tomé, Chile, Ciudad de T...	570	
67	Malaquías Concha 100, Tome, Tomé, Chile, Ciuda...	694	
124	Pasaje 1 3600, Talcahuano, Chile, Talcahuano, ...	217	
163	Gabriel Toro 198, Talcahuano, Chile, Carriel, ...	400	
166	Andrés Bello 361, 4260332 Concepcion, Talcahua...	500	
251	Carriel, Talcahuano, Biobío	740	
256	Supermercado Olimpia / Cruce Higuieras, Carrie...	800	
365	Avenida Las Margaritas, San Pedro De La Paz, C...	650	
372	Se Vende Moderna Casa ,semi Nueva, Huertos Fam...	500	
459	San Pedro De La Paz, Huertos Familiares, San P...	649	
522	Huertos Familiares, San Pedro De La Paz, Huert...	650	
528	Andalue Sn, Andalué, San Pedro De La Paz, Biobío	377	
559	Se Vende Moderna Casa ,semi Nueva, Huertos Fam...	500	
654	Las Violetas 1800 - 2100, San Pedro De La Paz,...	680	
710	Futaleufú 1135, San Pedro De La Paz, Chile, An...	450	
1212	Génova 2681, Hualpén, Talcahuano, Bío Bío, Chi...	160	
1324	No disponible	382	
1463	Victoria/chorrillos, Barrio Universitario, Con...	1000	
1493	Pedro Osores De Ulloa, Concepción, Chile, Loma...	905	
1522	Diagonal Oriente, Concepción, Chile, Barrio Or...	363	
1531	Concepción, Biobío	578	
1536	Concepción, Chile, Centro de Concepción, Conce...	460	
1542	Sector Sanatorio Aleman, Pedro de Valdivia, Co...	585	
1549	Barrio Universitario, Concepción, Chile, Barri...	552	
1552	Barrio Universitario, Concepción, Barrio Unive...	1000	
1556	Lautaro 857, Concepción, Chile, Barrio Oriente...	366	
1559	Pasaje 1 4872, Barrio Poniente, Concepción, Bi...	460	
1570	Los Olmos 1180, Concepción, Chile, Barrio Univ...	750	
1578	Cercano Parque Ecuador, Centro de Concepción, ...	317	
1608	Cochrane 1343, Concepción, Bío Bío, Chile, Bar...	304	
1626	Manuel Bulnes 1000, Barrio Poniente, Concepció...	617	
1638	Serrano 363, Concepción, Chile, Centro de Conc...	700	
1640	Serrano 363, Concepción, Chile, Centro de Conc...	700	
1652	Diagonal Pedro Aguirre Cerda, Centro de Concep...	534	

1673	Anibal Pinto 20, Concepción, Chile, Centro de ...	309
1685	Sector Barrio Universitario, Barrio Universita...	750
1694	Luis Urrutia Manzano 478, 4070333 Concepción, ...	532
1707	Los Lirios, Concepción, Chile, Collao, Concepc...	162
1722	Angol 1152, Barrio Poniente, Concepción, Biobío	617
1725	Isabel Riquelme, Collao, Concepción, Biobío	178
1762	Barrio Universitario, Concepción, Biobío	1000
1784	Avenida San Juan Bosco, Concepción, Chile, Bar...	322
1791	Tucapel, Concepción, Chile, Barrio Poniente, C...	222
1798	Maipú, Concepción, Chile, Centro de Concepción...	363
1801	Obispo Hipólito Salas 1094, Concepción, Chile,...	451
1803	Concepcion, Barrio Universitario, Concepción, ...	1000
1811	Lientur 1264, Barrio Oriente - Las Tres Pascua...	450
1828	Las Canchas 3336, Concepción, Chile, Pedro de ...	500
1829	Brasil 944, Bajo Tucapel, Concepción, Biobío	200
1840	Las Lomas San Andres, Lomas de San Sebastián, ...	975
2408	Centro de Concepción, Concepción, Biobío	395
2809	Antumalal, Lonco, Chiguayante, Biobío	610
2822	Estadio Español, Schaub, Chiguayante, Biobío	453
2852	Walter Schaub, Chiguayante, Chile, Schaub, Chi...	729
2853	Walter Schaub, Chiguayante, Chile, Schaub, Chi...	729
2862	Avenida Inglesa 560, Lonco, Chiguayante, Biobío	340
2897	Lonco Parque, Lonco, Chiguayante, Biobío	1400
2912	Ramón Harriet 120, Concepción, Chile, Lonco, C...	969

	Dormitorios	Baños
57	9	4
60	9	5
67	8	4
124	7	2
163	7	2
166	7	2
251	7	3
256	8	3
365	10	7
372	8	6
459	10	6
522	10	7
528	10	3
559	8	6
654	10	7
710	8	4
1212	7	1
1324	7	2
1463	9	5
1493	8	5
1522	11	5
1531	9	4
1536	7	5
1542	9	4
1549	9	4
1552	10	8
1556	7	3
1559	7	5
1570	8	3
1578	18	15
1608	8	7
1626	7	3
1638	10	6
1640	10	6
1652	15	4
1673	8	5
1685	8	3
1694	7	4
1707	7	3
1722	7	5
1725	10	4
1762	11	10
1784	8	3
1791	7	2
1798	11	5
1801	9	6

1803	9	8
1811	8	2
1828	12	8
1829	7	4
1840	7	6
2408	11	6
2809	7	6
2822	7	3
2852	7	5
2853	7	5
2862	8	5
2897	7	6
2912	7	5

Al observar todas las características de estos datos se decide mantenerlos ya que son propiedades de un tamaño considerable para mantener todas estas habitaciones y muchas de ellas están en barrios universitarios por lo que pueden estar asociadas al arriendo de piezas para universitarios por lo que no se considerarán como errores de tipeo.

In [23]:

```
df
```

Out[23]:

	Comuna	Propiedad	Precio	Direccion	Tamaño en m2	Dormitorios	Baños
0	Tomé	Casa	221248916	Pingüeral, Tomé, Biobío	559	4	2
1	Tomé	Casa	291797402	Avenida Pingüeral, Tomé, Chile, Pingüeral, Tomé,...	1068	4	4
2	Tomé	Casa	299923405	Bella Casa Con Piscina 3d 2b T350 M2 , Tomé, T...	350	3	2
3	Tomé	Casa	254861021	Avenida Pingüeral, Tomé, Chile, Pingüeral, Tomé,...	560	4	3
4	Tomé	Casa	1244756006	Avenida Pingüeral, Tomé, Chile, Pingüeral, Tomé,...	875	6	4
...	...	...	...	...	...	...	...
3109	Chiguayante	Departamento	60000000	Esperanza 507, Chiguayante, Chile, La Leonera,...	53	3	1
3111	Chiguayante	Departamento	184681900	La Leonera, Chiguayante, Biobío	85	2	2
3112	Chiguayante	Departamento	74000000	Esperanza 507, Chiguayante, Chile, La Leonera,...	59	3	1
3113	Chiguayante	Departamento	350000000	Sanhueza 55, Concepción, Chile, Lonco, Chiguay...	145	4	3
3114	Chiguayante	Departamento	151439158	Fresia 90, Chiguayante, Chile, Altos de Chigua...	90	3	2

2962 rows x 7 columns

In [24]:

```
#Se agrega una columna ID al dataframe
df['ID'] = df.index
```

In [25]:

```
#Se observan los datos por ultima vez
df.info()
```

<class 'pandas.core.frame.DataFrame'>
Int64Index: 2962 entries, 0 to 3114
Data columns (total 8 columns):

```
Data columns (total 8 columns):  
#      Column      Non-Null Count  Dtype  
---  -  
0     Comuna        2962 non-null    object  
1     Propiedad      2962 non-null    object  
2     Precio         2962 non-null    int32  
3     Direccion      2962 non-null    object  
4     Tamaño en m2   2962 non-null    int64  
5     Dormitorios    2962 non-null    int64  
6     Baños           2962 non-null    int64  
7     ID             2962 non-null    int64
```

```
dtypes: int32(1), int64(4), object(3)
```

```
memory usage: 196.7+ KB
```

```
In [26]:
```

```
#SE EXPORTA UN NUEVO ARCHIVO
```

```
df.to_csv("pgcf.csv", index=False)
```