Home Rentals in San Francisco Report

Introduction

Scenario

Every one of us has heard of San Francisco. Over the last few decades, this beautiful city has grown so much in popularity that millions of people come to feel the vibe and see famous locations every year. These include the world's steepest street (Lombard Street) and the famous prison (Alcatraz). Visitors come to eat delicious food in one of many splendid restaurants, like the Bubba Gump Shrimp Company or numerous Tex-Mex bakeries or restaurants, where you can get the best tacos and burritos.

San Francisco serves as a go-to location for digital nomads lured to the city by the biggest tech companies in the world (Meta, Google, and Apple, to name a few), who have set up their headquarters in nearby Silicon Valley.

This whole phenomenon has its dark side. Due to the growing popularity, the real estate market in San Francisco went up at such a crazy rate that more and more people were forced to abandon their American dream of a nice house in the city centre and either look for a home elsewhere or simply become homele

Dataset

In this study, we'll analyse the house rentals data from Craigslist and prepare a presentation showing the real estate situation in San Francisco.isco.

```python
#Import library
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as snb
# Import the dataset
HomeRental = pd.read_csv(r"C:\Users\tuana\OneDrive\Documents\Thu Hue Duong\DATA ANALYST\ANALYSIS REPORT\sf_clean.csv'
# An overview of the dataset
print(HomeRental.shape)# The number of rows and columns
print(HomeRental.head(10))# Show the first 10 rows in the data
```
In [6]:

```python
print(HomeRental.describe())# Show the statistical parameters on each column
print(HomeRental.info())# Show the properties of the data
```

```
(989, 9)
    price    sqft  beds  bath         laundry            pets housing_type  \
0    6800  1600.0   2.0   2.0     (a) in-unit    (d) no pets    (c) multi
1    3500   550.0   1.0   1.0     (a) in-unit      (a) both    (c) multi
2    5100  1300.0   2.0   1.0     (a) in-unit      (a) both    (c) multi
3    9000  3500.0   3.0   2.5     (a) in-unit    (d) no pets    (c) multi
4    3100   561.0   1.0   1.0  (c) no laundry      (a) both    (c) multi
5    3800   800.0   2.0   1.0     (b) on-site      (c) cats    (c) multi
6    3100   750.0   1.0   1.0     (a) in-unit    (d) no pets    (c) multi
7    3000   650.0   1.0   1.0     (b) on-site      (a) both    (c) multi
8    3000   650.0   1.0   1.0     (b) on-site      (a) both    (c) multi
9    3200   650.0   1.0   1.0     (a) in-unit      (c) cats    (c) multi

           parking  hood_district
0     (b) protected            7.0
1     (b) protected            7.0
2    (d) no parking            7.0
3     (b) protected            7.0
4    (d) no parking            7.0
5     (b) protected            9.0
6     (b) protected            8.0
7    (d) no parking            7.0
8     (b) protected            7.0
9     (b) protected            9.0
             price         sqft         beds         bath  hood_district
count   989.000000   989.000000   989.000000   989.000000     989.000000
mean   3595.035389   976.765420     1.679474     1.390293       7.052578
std    1546.222670   474.629798     1.076710     0.562714       2.404716
min     750.000000   150.000000     0.000000     1.000000       1.000000
25%    2650.000000   650.000000     1.000000     1.000000       6.000000
50%    3300.000000   900.000000     2.000000     1.000000       8.000000
75%    4242.000000  1200.000000     2.000000     2.000000       9.000000
max   19000.000000  3500.000000     6.000000     4.000000      10.000000
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 989 entries, 0 to 988
Data columns (total 9 columns):
 #   Column          Non-Null Count  Dtype
---  ------          --------------  -----
 0   price           989 non-null    int64
 1   sqft            989 non-null    float64
 2   beds            989 non-null    float64
 3   bath            989 non-null    float64
```

```
4    laundry        989 non-null    object
5    pets           989 non-null    object
6    housing_type   989 non-null    object
7    parking        989 non-null    object
8    hood_district  989 non-null    float64
dtypes: float64(4), int64(1), object(4)
memory usage: 69.7+ KB
```

Analyzing data

The popular district for rent in San Francisco

In [61]:
```python
# NORental=HomeRental.groupby(['hood_district']).count()
# print(NORental)
```

In [63]:
```python
#Show the histogram of hood_district column to identify the most popular rental district.
plt.hist(HomeRental.hood_district)
plt.title('The number of home rentals by districts')
plt.xlabel('Hood district')
plt.ylabel('The number of rentals')
# Finding the mode in a column by mode function
#HomeRental.hood_district.mode()
```

Out[63]: Text(0, 0.5, 'The number of rentals')

## The number of home rentals by districts
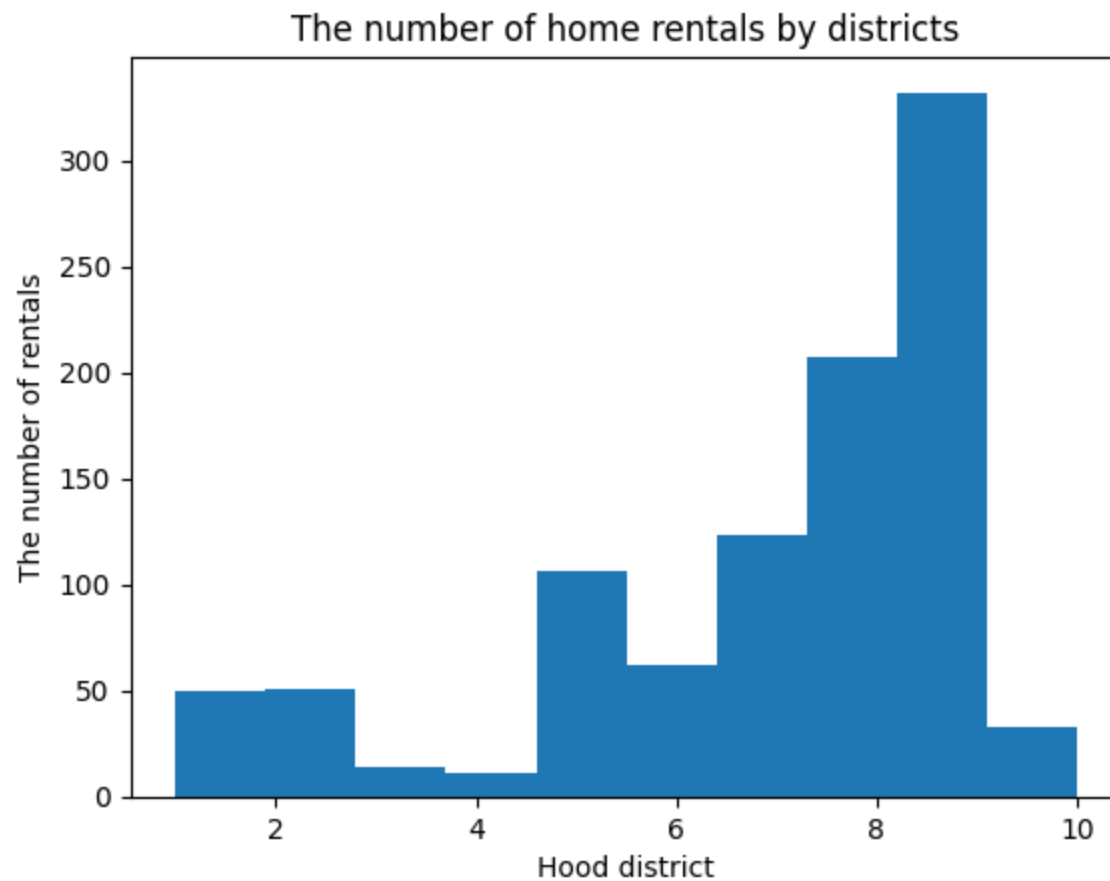
It is noticeable that the hood district number 9 is the most popular rental place in San Francisco. There are about 350 apartments or houses for rental in district 9, which is 1/3 the total of house rentals in San Francisco.

The most frequently offered number of bedrooms

```
In [64]:  # Show the histogram of beds column
          plt.hist(HomeRental.beds)
          plt.title('The number of home rentals by beds')
          plt.xlabel('The number of bedrooms')
          plt.ylabel('The number of rentals')
```
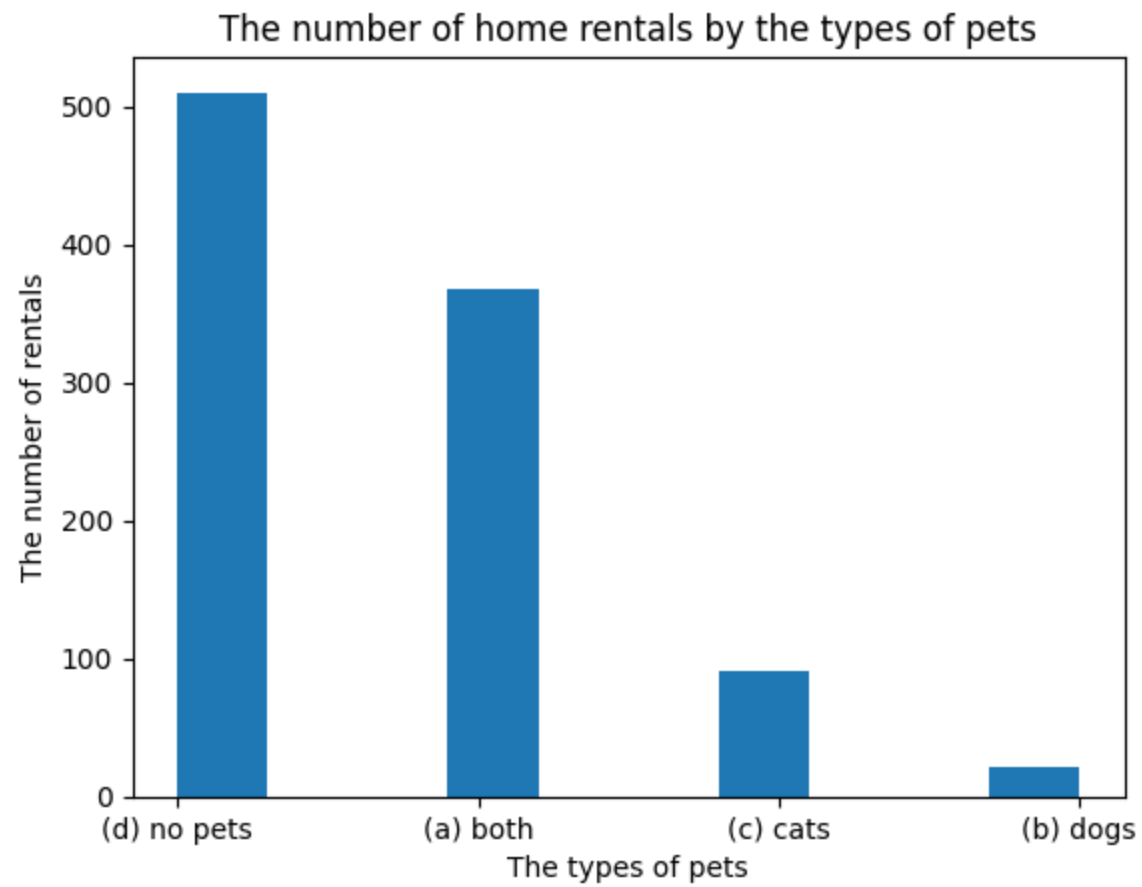
```
Out[64]:  Text(0, 0.5, 'The number of rentals')
```

The number of home rentals by beds
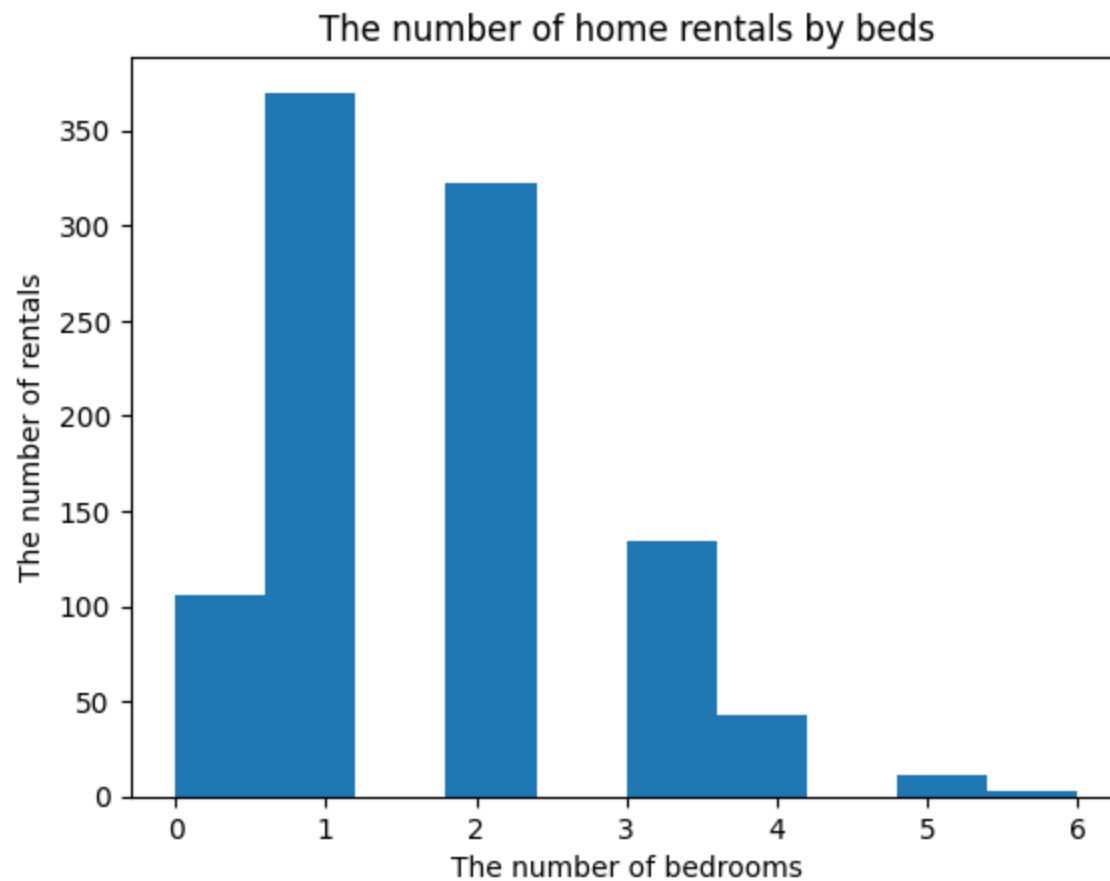
We see that the 1 bedroom apartment is the most common in real estate market in San Francisco with axproximate 370 offers.

Find an apartment that allows having animals or not

```python
In [67]:  # Show the histogram of pets column
          plt.hist(HomeRental.pets)
          plt.title('The number of home rentals by the types of pets')
          plt.xlabel('The types of pets')
          plt.ylabel('The number of rentals')
```

```
Out[67]:  Text(0, 0.5, 'The number of rentals')
```

The number of home rentals by the types of pets

In [ ]: