

```

In [7]: import matplotlib.pyplot as plt
import pandas as pd
import csv
import numpy as np
import matplotlib

In [8]: with open('district9_rent.csv', newline='', encoding='utf-8') as f:
reader = csv.reader(f, delimiter=',')
for row in reader:
print(row)

30131,dunglinht@yahoo.com,,Xã Lộ Hà Nội,,,Cho thuê,,55.0'
['1,123,53,SG,61,Hồ Chí Minh,Quận 9,8008851,130.0,2.0,1500.0,10,8,5844,-1,-1,KXĐ,Không thuộc dự á
n,-1,"Phường Trường Thạnh, Quận 9, Hồ Chí Minh",,,,10.810220663351885,106.81570959161013,"Cho thu
ê kho, nhà xưởng, đất",2015-08-10,2015-10-13,Cho thuê kho xưởng 1500m2 2500m2 3500m2 phường Trường
Thạnh quận Gân khu công nghệ cao Xưởng mới xây rất đẹp đầy đủ trạm điện pha nước bãi đậu xe rộng Gi
ao thông thuận tiện đường xe cont giá 130tr/tháng LH 0908130131,"Cho thuê kho xưởng 1500m2, 2500m2,
3500m2 phường Trường Thạnh, quận 9. Gân khu công nghệ cao. Xưởng mới xây rất đẹp, đầy đủ trạm điện

```

```

In [9]: result = pd.read_csv('district9_rent.csv')
print(result)

   Unnamed: 0  index  CateId  CityCode  DistrictId  CityName \
0           0      28      53        SG          61  Hồ Chí Minh
1           1      123      53        SG          61  Hồ Chí Minh
2           2      207      326       SG          61  Hồ Chí Minh
3           3      238       52        SG          61  Hồ Chí Minh
4           4      243       55        SG          61  Hồ Chí Minh
...         ...    ...    ...    ...    ...    ...
16313      16313  357753       52        SG          61  Hồ Chí Minh
16314      16314  357758      326        SG          61  Hồ Chí Minh
16315      16315  357843       52        SG          61  Hồ Chí Minh
16316      16316  357883      326        SG          61  Hồ Chí Minh
16317      16317  357899      326        SG          61  Hồ Chí Minh

   DistrictName  ProductId  Price  UnitPrice  ... \
0      Quận 9      8006014    55.0        2.0  ...
1      Quận 9      8008851   130.0        2.0  ...
2      Quận 9      8012743     4.0        2.0  ...
3      Quận 9      8017058     3.5        2.0  ...

```

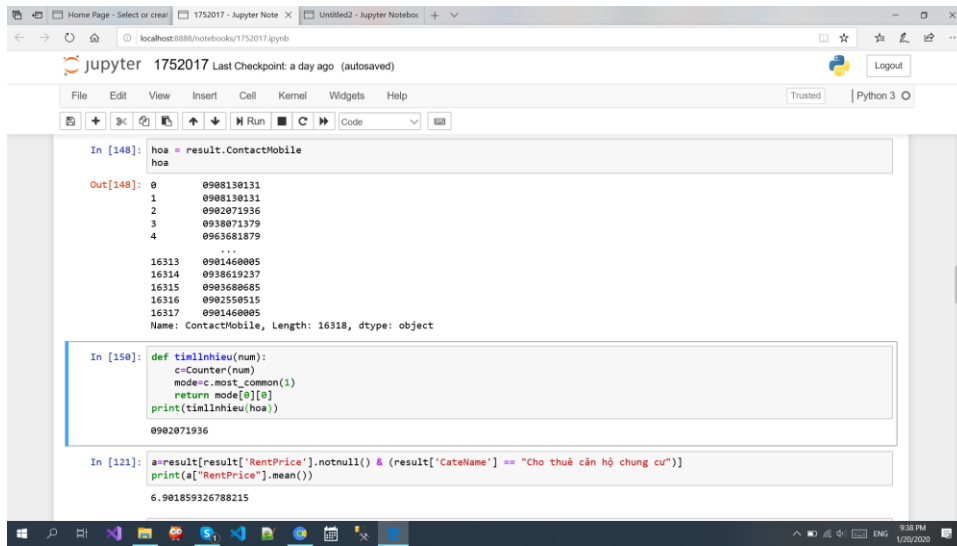
```

In [10]: result

Out[10]:
   Unnamed: 0  index  CateId  CityCode  DistrictId  CityName  DistrictName  ProductId  Price  UnitPrice  ...  ContactAddress:
0           0      28      53        SG          61  Hồ Chí Minh  Quận 9      8006014    55.0        2.0  ...  Nat
1           1      123      53        SG          61  Hồ Chí Minh  Quận 9      8008851   130.0        2.0  ...  Nat
2           2      207      326       SG          61  Hồ Chí Minh  Quận 9      8012743     4.0        2.0  ...  ho chi mini
3           3      238       52        SG          61  Hồ Chí Minh  Quận 9      8017058     3.5        2.0  ...  Nat
4           4      243       55        SG          61  Hồ Chí Minh  Quận 9      8018051    NaN        NaN  ...  89/1/49/11
                                     đường 8 kpi
                                     p.Tầng Nh...
                                     Tp.HCM
...         ...    ...    ...    ...    ...    ...
16313      16313  357753       52        SG          61  Hồ Chí Minh  Quận 9     19982945   10.0        2.0  ...  Sơn Hòa, Hương
                                     Sơn, Hà Tĩnh

```

1, 2



The screenshot shows a Jupyter Notebook window with the following content:

```
In [148]: hoa = result.ContactMobile
          hoa

Out[148]: 0      0908130131
          1      0908130131
          2      0902071936
          3      0938071379
          4      0963681879
          ...
          16313  0901460005
          16314  0938619237
          16315  0903680685
          16316  0902590515
          16317  0901460005
          Name: ContactMobile, Length: 16318, dtype: object

In [150]: def timltnhieu(num):
          c=Counter(num)
          mode=c.most_common(1)
          return mode[0][0]
          print(timltnhieu(hoa))

          0902071936

In [121]: a=result[result['RentPrice'].notnull() & (result['CateName'] == "Cho thuê căn hộ chung cư")]
          print(a["RentPrice"].mean())

          6.901859326788215
```

3, from matplotlib import pyplot as plt

Sử dụng thư viện matplotlib để vẽ biểu đồ

---> Lấy ra giá trị trung bình từng năm.

Nam= [2015,2016,2017,2018,2019]

Gia_tb = [20.1, 21.2, 17.9, 23.0, 24.6] # ví dụ

xs = [i + 0.5 for i, _ in enumerate(Nam)]

plt.bar(xs, Gia_tb)

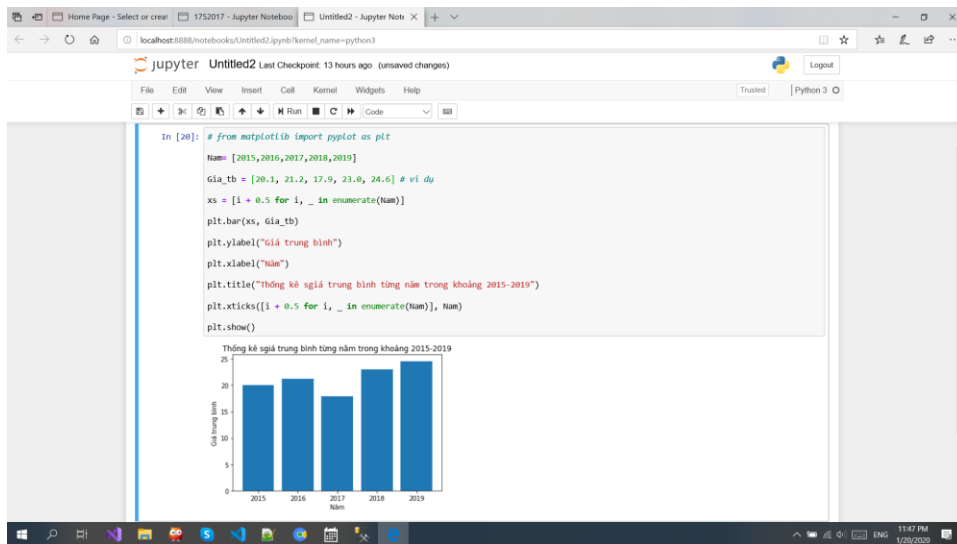
plt.ylabel("Giá trung bình")

plt.xlabel("Năm")

plt.title("Thống kê giá trung bình từng năm trong khoảng 2015-2019")

plt.xticks([i + 0.5 for i, _ in enumerate(Nam)], Nam)

plt.show()



4,

Jupyter Notebook interface showing a Python script and its output.

```
In [28]: projName=result.groupby('ProjectName');
         type(projName)
         pd.core.groupby.DataFrameGroupBy
         len(projName.groups)

Out[28]: 84

In [29]: projName.groups
```

```
dtype='int64', length=11/),
'Mega Residence': Int64Index([ 142, 145, 150, 162, 187, 201, 234, 245, 247,
                               249, 300, 311, 333, 480, 541, 752, 808, 822,
                               857, 882, 1019, 1093, 1123, 1151, 1153, 1156, 1158,
                               1352, 1428, 1485, 1495, 1883, 1964, 2016, 2059, 2112,
                               2220, 2467, 2692, 3708, 3764, 4069, 4585, 4830, 4846,
                               4940, 5021, 5132, 5506, 5548, 5961, 6131, 6158, 6181,
                               6209, 6220, 6294, 6369, 6392, 6633, 6691, 7309, 7373,
                               7441, 7531, 7718, 7971, 8885, 9915, 10267, 10860, 11988,
                               12450, 12451, 13691, 13860, 13924, 14342, 15042, 15212, 15694],
                               dtype='int64'),
'Mega Ruby': Int64Index([ 283, 289, 299, 722, 727, 817, 866, 945, 970,
                           972,
```

```
1752017 - Jupyter Note
Untitled2 - Jupyter Notebook
localhost:8888/notebooks/1752017.ipynb

In [124]: projName=result.groupby('ProjectName');
          type(projName)
          pd.core.groupby.DataFrameGroupBy
          len(projName.groups)

Out[124]: 84

In [165]: prjn=projName['RentPrice'].count();
          prjn

Out[165]: ProjectName
9 View Apartment      8
Biệt thự Cielo Island 1
Cao ốc Bình Minh     29
Cao ốc Phố Đông - Hoa Sen 73
Centana Điền Phúc Thành 1
..
Valora Fuji          13
Villa Park           43
Vincity quận 9        0
Việt Nhân Villa Riverside 2
Vườn Lan             1
Name: RentPrice, Length: 84, dtype: int64
```

```
1752017 - Jupyter Note
Untitled2 - Jupyter Notebook
localhost:8888/notebooks/1752017.ipynb

In [166]: banggia=result[['ProjectName','RentPrice']]
          banggia

Out[166]:
```

	ProjectName	RentPrice
0	Không thuộc dự án	55.0
1	Không thuộc dự án	130.0
2	Nhân Phú	4.0
3	Không thuộc dự án	3.5
4	Không thuộc dự án	NaN
...
16313	Không thuộc dự án	10.0
16314	Him Lam Phú An	10.0
16315	Không thuộc dự án	7.0
16316	Jamila Khang Điền	8.0
16317	Lucky Dragon	7.5

16318 rows x 2 columns

```
1752017 - Jupyter Note
Untitled2 - Jupyter Notebook
localhost:8888/notebooks/1752017.ipynb

In [172]: canh=banggia[banggia['RentPrice'].notnull() & (result['ProjectName'] != "Không thuộc dự án")]
          canh

Out[172]:
```

	ProjectName	RentPrice
2	Nhân Phú	4.00
9	KDC Khang Điền - Intresco	16.00
15	Villa Park	30.00
17	Villa Park	55.58
18	Khu dân cư Nam Long	12.00
...
16310	The Art	7.50
16312	Flora Anh Đào	6.00
16314	Him Lam Phú An	10.00
16316	Jamila Khang Điền	8.00
16317	Lucky Dragon	7.50

8766 rows x 2 columns

