

北京邮电大学

Python 程序设计实验报告 ——租房数据分析



姓 名	<u>董奕柳</u>
班 级	<u>2021211304</u>
学 号	<u>2021210868</u>
任课教师	<u>杨亚</u>

2023 年 12 月

目录

- 1 要求3
- 2 过程及核心代码3
 - 2.1 爬虫3
 - 2.1.1 说明.....3
 - 2.1.2 代码及结果.....3
 - 2.2 数据处理8
 - 2.2.1 说明.....8
 - 2.2.2 代码及结果.....8
 - 2.3 数据可视化..... 12
 - 2.3.1 说明..... 12
 - 2.3.2 代码..... 12
- 3 可视化结果及分析 21
 - 3.1 总价和平米均价 21
 - 3.2 不同布局总价 22
 - 3.3 各城市不同板块均价 23
 - 3.4 不同朝向对租金的影响..... 26
 - 3.5 2023 与 2022 对比..... 29
 - 3.6 工资与租房价格 29
- 4 报告总结..... 30

1 要求

1. 要求抓取链家官网北上广深 4 个一线城市，再加上一个离你家乡最近的一个非一线城市/或者你最感兴趣的一个城市的租房数据。
2. 应获取每个城市的全部租房数据（一线城市的数据量应该在万的数量级）。
3. 比较 5 个城市的总体房租情况，包含租金的均价、最高价、最低价、中位数等信息，单位面积租金（元/平米）的均价、最高价、最低价、中位数等信息。采用合适的图或表形式进行展示。
4. 比较 5 个城市一居、二居、三居的情况，包含均价、最高价、最低价、中位数等信息，采用合适的图或表形式进行展示。
5. 计算和分析每个城市不同板块的均价情况，并采用合适的图或表形式进行展示。
6. 比较各个城市不同朝向的单位面积租金分布情况，采用合适的图或表形式进行展示。哪个方向最高，哪个方向最低？各个城市是否一致？如果不一致，你认为原因是什么？
7. 查询各个城市的平均工资，分析并展示其和单位面积租金分布的关系。比较一下在哪个城市租房的负担最重？
8. 与 2022 年的租房数据进行对比（只比较北上广深 4 个城市，原始数据会给出），总结你观察到的变化情况，并用图、表、文字等支撑你得到的结论。

2 过程及核心代码

2.1 爬虫

2.1.1 说明

第一次爬时，直接爬了主页，但超过 100 页后是重复数据，因此要分治，到板块里读。板块则需要从每个区里爬取，区需要从城市里爬取。因此有三步爬虫。

2.1.2 代码及结果

先爬区及链接。

```

spiderall > 1spider_link_district > lianjia > spiders > spider_beijing.py > LianjiaSpider > __init__
1  # spider.py
2
3  import scrapy
4
5  from lianjia.items import LianjiaItem
6
7
8  class LianjiaSpider(scrapy.Spider):
9      name = "lianjia_beijing"
10     allowed_domains = ["bj.lianjia.com"]
11     start_urls = ["https://bj.lianjia.com/zufang/"]
12
13     def __init__(self, **kwargs):
14         super().__init__(**kwargs)
15         self.download_delay = 5
16
17     def parse(self, response, **kwargs):
18         print(response.body)
19         item = LianjiaItem()
20
21         info_list = response.xpath('//li[@data-type="district"]')
22
23         print("info_list:", info_list)
24         for info in info_list:
25             item["name"] = info.xpath("./a/text()").get()
26             item["link_district"] = info.xpath("./a/@href").get()
27
28             yield item
29

```

```

spiderall > 1spider_link_district > {} lianjia_shanghai.json > ...
1  {"name": "不限", "link_district": "/zufang/"}
2  {"name": "静安", "link_district": "/zufang/jingan/"}
3  {"name": "徐汇", "link_district": "/zufang/xuhui/"}
4  {"name": "黄浦", "link_district": "/zufang/huangpu/"}
5  {"name": "长宁", "link_district": "/zufang/changning/"}
6  {"name": "普陀", "link_district": "/zufang/putuo/"}
7  {"name": "浦东", "link_district": "/zufang/pudong/"}
8  {"name": "宝山", "link_district": "/zufang/baoshan/"}
9  {"name": "虹口", "link_district": "/zufang/hongkou/"}
10 {"name": "杨浦", "link_district": "/zufang/yangpu/"}
11 {"name": "闵行", "link_district": "/zufang/minhang/"}
12 {"name": "金山", "link_district": "/zufang/jinshan/"}
13 {"name": "嘉定", "link_district": "/zufang/jiading/"}
14 {"name": "崇明", "link_district": "/zufang/chongming/"}
15 {"name": "奉贤", "link_district": "/zufang/fengxian/"}
16 {"name": "松江", "link_district": "/zufang/songjiang/"}
17 {"name": "青浦", "link_district": "/zufang/qingpu/"}
18

```

再到每个区里爬板块。

```

3 import scrapy
4
5 from lianjia.items import LianjiaItem
6
7 import json
8
9
10 class LianjiaSpider(scrapy.Spider):
11     name = "lianjia_beijing"
12     allowed_domains = ["bj.lianjia.com"]
13
14     with open(
15         "../spider_link_district/lianjia_beijing.json", "r", encoding="utf-8"
16     ) as f:
17         url_list = []
18         for line in f:
19             url_list.append(json.loads(line))
20
21     start_urls = []
22     for url in url_list:
23         url = "https://bj.lianjia.com" + url["link_district"]
24         print(url)
25         start_urls.append(url)
26
27     def __init__(self, **kwargs):
28         super().__init__(**kwargs)
29         self.download_delay = 5 # 设置下载延迟,防止被封
30
31     def parse(self, response, **kwargs):
32         item = LianjiaItem()
33
34         info_list = response.xpath('//li[@class="filter__item--level3 "]')
35
36         print("info_list:", info_list)
37         for info in info_list:
38             item["name"] = info.xpath("./a/text()").get()
39             item["link_district"] = info.xpath("./a/@href").get()
40
41         yield item
42

```

```

spiderall > 2spider_link_bankuai > {} lianjia_beijing.json > ...
1 {"name": "安定门", "link_district": "/zufang/andingmen/"}
2 {"name": "安贞", "link_district": "/zufang/anzhen1/"}
3 {"name": "朝阳门外", "link_district": "/zufang/chaoyangmenwai/"}
4 {"name": "朝阳门内", "link_district": "/zufang/chaoyangmennei/"}
5 {"name": "崇文门", "link_district": "/zufang/chongwenmen/"}
6 {"name": "东单", "link_district": "/zufang/dongdan/"}
7 {"name": "东直门", "link_district": "/zufang/dongzhimen/"}
8 {"name": "地安门", "link_district": "/zufang/dianmen/"}
9 {"name": "东花市", "link_district": "/zufang/donghuashi/"}
10 {"name": "东四", "link_district": "/zufang/dongsil/"}
11 {"name": "灯市口", "link_district": "/zufang/dengshikou/"}
12 {"name": "广渠门", "link_district": "/zufang/guangqumen/"}
13 {"name": "工体", "link_district": "/zufang/gongti/"}
14 {"name": "和平里", "link_district": "/zufang/hepingli/"}
15 {"name": "交道口", "link_district": "/zufang/jiaodaokou/"}
16 {"name": "建国门外", "link_district": "/zufang/jianguomenwai/"}
17 {"name": "建国门内", "link_district": "/zufang/jianguomennei/"}
18 {"name": "金宝街", "link_district": "/zufang/jinbaojie/"}
19 {"name": "六铺炕", "link_district": "/zufang/liupukang/"}
20 {"name": "蒲黄榆", "link_district": "/zufang/puhuangyu/"}
21 {"name": "前门", "link_district": "/zufang/qianmen/"}
22 {"name": "陶然亭", "link_district": "/zufang/taoranting1/"}
23 {"name": "天坛", "link_district": "/zufang/tiantan/"}
24 {"name": "西罗园", "link_district": "/zufang/xiluoyuan/"}
25 {"name": "西单", "link_district": "/zufang/xidan/"}
26 {"name": "永定门", "link_district": "/zufang/yongdingmen/"}
27 {"name": "洋桥", "link_district": "/zufang/yangqiao1/"}
28 {"name": "左安门", "link_district": "/zufang/zuonanmen1/"}
29 {"name": "白纸坊", "link_district": "/zufang/baizhifang1/"}
30 {"name": "菜户营", "link_district": "/zufang/caihuying/"}
31 {"name": "长椿街", "link_district": "/zufang/changchunjie/"}
32 {"name": "车公庄", "link_district": "/zufang/chegongzhuang1/"}
33 {"name": "地安门", "link_district": "/zufang/dianmen/"}
34 {"name": "德胜门", "link_district": "/zufang/deshengmen/"}

```

最后爬租房信息。

```
# spider.py

import scrapy

from lianjia.items import LianjiaItem

import json

class LianjiaSpider(scrapy.Spider):
    name = "lianjia_shanghai"
    allowed_domains = ["sh.lianjia.com"]
    with open(
        "../2spider_link_bankuai/lianjia_shanghai_processed.json", "r",
encoding="utf-8"
    ) as f:
        url_list = [json.loads(line) for line in f]

        start_urls = ["https://sh.lianjia.com" +
url_list[0]["link_district"]]

        page_index = 1

    def __init__(self, **kwargs):
        super().__init__(**kwargs)
        self.download_delay = 5

    def parse(self, response, **kwargs):
        item = LianjiaItem()

        house_total = response.xpath("//span[@class='content__title--
hl']/text()").get()
        house_total = int(house_total)
        print("house_total:", house_total)
        page_total = house_total // 30 + 1
        print("page_total:", page_total)

        info_list = response.xpath('//div[@class="content__list--item--
main"]')
        for info in info_list:
            item["name"] = info.xpath(
                './p[@class="content__list--item--title"]/a/text()'
            ).get()
```

```

        if not item["name"]:
            item["name"] = info.xpath(
                './p[@class="content__list--item--title
twoline"]/a/text()'
            ).get()
        if item["name"]:
            item["name"] = item["name"].strip()
            # print(item["name"])

        item["district"] = info.xpath(
            './p[@class="content__list--item--des"]/a[2]/text()'
        ).get()

        item["total_price"] = info.xpath(
            './span[@class="content__list--item-price"]/em/text()'
        ).get()

        item["area_direction_layout"] = info.xpath(
            './p[@class="content__list--item--des"]/text()'
        ).extract()
        if item["area_direction_layout"]:
            item["area_direction_layout"] = [
                i.strip() for i in item["area_direction_layout"] if
i.strip()
            ]
        yield item

    print("self.page_index:", self.page_index)
    print("page_total:", page_total)

    if self.page_index >= page_total or page_total == 0:
        self.page_index = 0
        self.url_list = self.url_list[1:]

    self.page_index += 1

    if self.url_list:
        url = (
            "https://sh.lianjia.com"
            + self.url_list[0]["link_district"]
            + "pg"
            + str(self.page_index)
        )
        yield scrapy.Request(url, callback=self.parse)

```

```
spiderall > 3_spider > [] lianjia.beijingjion > ...
1 {'name': '整租·安外东河沿 1室1厅 南', 'district': '安定门', 'total_price': '6000', 'area_direction_layout': ['-', '-', '39.00m²', '南', '1室1厅1卫']}
2 {'name': '整租·安德路5号院 2室1厅 北', 'district': '安定门', 'total_price': '7330', 'area_direction_layout': ['舒适', '-', '-', '40.21m²', '北', '2室1厅1卫']}
3 {'name': '整租·安外西河沿路1号院 3室1厅 南', 'district': '安定门', 'total_price': '17000', 'area_direction_layout': ['-', '-', '113.00m²', '南', '3室1厅1卫']}
4 {'name': '整租·共享际公寓 金台·共享际 套间 1室1厅', 'district': null, 'total_price': '18000', 'area_direction_layout': ['42.86m²', '1间在租', '1室1厅1卫']}
5 {'name': '整租·安外西河沿 1室1厅 南', 'district': '安定门', 'total_price': '5900', 'area_direction_layout': ['-', '-', '43.39m²', '南', '1室1厅1卫']}
6 {'name': '整租·城家公寓 马甸黄寺大街店 1个月起租 可月付 开间', 'district': null, 'total_price': '5100-5600', 'area_direction_layout': ['21.71-23.86m²', '8间在租', '1室0厅1卫']}
7 {'name': '整租·六铺炕 3室1厅 南/北', 'district': '安定门', 'total_price': '13000', 'area_direction_layout': ['-', '-', '73.00m²', '南 北', '3室1厅1卫']}
8 {'name': '整租·共享际公寓 金台·共享际 开间 开间', 'district': null, 'total_price': '7500-13000', 'area_direction_layout': ['20.98-24.82m²', '5间在租', '1室0厅1卫']}
9 {'name': '整租·上龙西里 2室1厅 南', 'district': '安定门', 'total_price': '7000', 'area_direction_layout': ['-', '-', '61.00m²', '南', '2室1厅1卫']}
10 {'name': '整租·城家公寓 马甸黄寺大街店 精装直租 可月付 开间', 'district': null, 'total_price': '4050-5150', 'area_direction_layout': ['22.06-23.86m²', '2间在租', '1室0厅1卫']}
11 {'name': '整租·安德路47号院 2室0厅 南/北', 'district': '安定门', 'total_price': '7200', 'area_direction_layout': ['-', '-', '56.65m²', '南 北', '3室0厅1卫']}
12 {'name': '整租·城家公寓 马甸黄寺大街店 限时特惠直租可做饭 开间', 'district': null, 'total_price': '5500-5700', 'area_direction_layout': ['22.86m²', '8间在租', '1室0厅1卫']}
13 {'name': '整租·西营房 1室0厅 北', 'district': '安定门', 'total_price': '5300', 'area_direction_layout': ['-', '-', '36.31m²', '北', '1室0厅1卫']}
14 {'name': '整租·城家公寓 马甸黄寺大街店 企业特惠 免费开票 开间', 'district': null, 'total_price': '5150', 'area_direction_layout': ['21.43m²', '1间在租', '1室0厅1卫']}
15 {'name': '整租·巷上高园 2室1厅 东/南', 'district': '安定门', 'total_price': '11600', 'area_direction_layout': ['-', '-', '99.27m²', '东 南', '2室1厅1卫']}
16 {'name': '整租·安外东河沿 1室1厅 南', 'district': '安定门', 'total_price': '5800', 'area_direction_layout': ['-', '-', '37.64m²', '南', '1室1厅1卫']}
17 {'name': '整租·青年湖东里 2室1厅 东/西', 'district': '安定门', 'total_price': '8000', 'area_direction_layout': ['-', '-', '64.20m²', '东 西', '2室1厅1卫']}
18 {'name': '整租·城家公寓 马甸黄寺大街店 直租可月付 可短租 开间', 'district': null, 'total_price': '5050', 'area_direction_layout': ['28.57m²', '1间在租', '1室0厅1卫']}
19 {'name': '整租·西营房 1室1厅 西', 'district': '安定门', 'total_price': '6700', 'area_direction_layout': ['-', '-', '55.95m²', '西', '1室1厅1卫']}
20 {'name': '整租·安德路47号院 1室1厅 南', 'district': '安定门', 'total_price': '5500', 'area_direction_layout': ['-', '-', '36.70m²', '南', '1室1厅1卫']}
21 {'name': '整租·安德路261号院 3室1厅 南/北', 'district': '安定门', 'total_price': '9800', 'area_direction_layout': ['-', '-', '66.67m²', '南 北', '3室1厅1卫']}
22 {'name': '整租·城家公寓 马甸黄寺大街店 近安华桥站 可月付 开间', 'district': null, 'total_price': '4950-5150', 'area_direction_layout': ['28.57m²', '2间在租', '1室0厅1卫']}
23 {'name': '整租·青年湖东里 1室1厅 南', 'district': '安定门', 'total_price': '6000', 'area_direction_layout': ['-', '-', '38.27m²', '南', '1室1厅1卫']}
24 {'name': '整租·城家公寓 马甸黄寺大街店 独厨独卫 精装直租 开间', 'district': null, 'total_price': '5050-5300', 'area_direction_layout': ['22.86m²', '2间在租', '1室0厅1卫']}
25 {'name': '整租·黄寺大街28号院 3室1厅 东/北', 'district': '安定门', 'total_price': '9500', 'area_direction_layout': ['-', '-', '81.79m²', '东 北', '3室1厅1卫']}
26 {'name': '整租·城家公寓 马甸黄寺大街店 可短租有折扣交通便利 开间', 'district': null, 'total_price': '5050-5300', 'area_direction_layout': ['22.86m²', '2间在租', '1室0厅1卫']}
27 {'name': '整租·永康胡同1号院 1室1厅 东/西', 'district': '安定门', 'total_price': '6300', 'area_direction_layout': ['-', '-', '47.00m²', '东 西', '1室1厅1卫']}
28 {'name': '整租·城家公寓 马甸黄寺大街店 近地铁精装直租特惠 开间', 'district': null, 'total_price': '4750-6050', 'area_direction_layout': ['23.43-30.71m²', '2间在租', '1室0厅1卫']}
29 {'name': '整租·花园胡同 3室1厅 东', 'district': '安定门', 'total_price': '8500', 'area_direction_layout': ['-', '-', '59.68m²', '东', '3室1厅1卫']}
30 {'name': '整租·城家公寓 马甸黄寺大街店 近地铁可短租有折扣 开间', 'district': null, 'total_price': '5100-5700', 'area_direction_layout': ['23.43-26.43m²', '4间在租', '1室0厅1卫']}
31 {'name': '整租·黄寺大街28号院 1室1厅 东/北', 'district': '安定门', 'total_price': '9500', 'area_direction_layout': ['-', '-', '81.79m²', '东 北', '3室1厅1卫']}
32 {'name': '整租·大广公寓 鼓楼安德里北街地铁站一店 酒店式公寓 开间', 'district': null, 'total_price': '4700', 'area_direction_layout': ['25.00m²', '1间在租', '1室0厅1卫']}
33 {'name': '整租·安德路47号院 1室1厅 南', 'district': '安定门', 'total_price': '6500', 'area_direction_layout': ['-', '-', '49.21m²', '南', '1室1厅1卫']}
34 {'name': '整租·大广公寓 鼓楼安德里北街地铁站一店 经济特价房 开间', 'district': null, 'total_price': '3000', 'area_direction_layout': ['18.00m²', '1间在租', '1室0厅1卫']}
35 {'name': '整租·西营房 1室1厅 西', 'district': '安定门', 'total_price': '9000', 'area_direction_layout': ['-', '-', '85.00m²', '西', '2室1厅1卫']}
36 {'name': '整租·城家公寓 马甸黄寺大街店 独厨独卫精装 直租 开间', 'district': null, 'total_price': '5150', 'area_direction_layout': ['23.43m²', '1间在租', '1室0厅1卫']}
37 {'name': '整租·永康胡同1号院 1室1厅 东/西', 'district': '安定门', 'total_price': '6300', 'area_direction_layout': ['-', '-', '47.00m²', '东 西', '1室1厅1卫']}
38 {'name': '整租·大广公寓 鼓楼安德里北街地铁站一店 loft从真两居室介绍 1室1厅', 'district': null, 'total_price': '4000', 'area_direction_layout': ['25.00m²', '1间在租', '1室1厅1卫']}
39 {'name': '整租·花园胡同 3室1厅 东', 'district': '安定门', 'total_price': '8500', 'area_direction_layout': ['-', '-', '59.68m²', '东', '3室1厅1卫']}
40 {'name': '整租·城家公寓 马甸黄寺大街店 精装直租 同时特惠 开间', 'district': null, 'total_price': '5000-5500', 'area_direction_layout': ['22.57-26.43m²', '6间在租', '1室0厅1卫']}
41 {'name': '整租·安德路47号院 3室1厅 南/北', 'district': '安定门', 'total_price': '8500', 'area_direction_layout': ['-', '-', '71.00m²', '南 北', '3室1厅1卫']}
42 {'name': '整租·城家公寓 马甸黄寺大街店 可短租有折扣出行便利 开间', 'district': null, 'total_price': '5500-5600', 'area_direction_layout': ['26.43m²', '4间在租', '1室0厅1卫']}
43 {'name': '整租·郎家胡同 1室1厅 南', 'district': '安定门', 'total_price': '4100', 'area_direction_layout': ['-', '-', '16.30m²', '南', '1室1厅1卫']}
44 {'name': '整租·大广公寓 鼓楼安德里北街地铁站一店 经济特价房 开间', 'district': null, 'total_price': '3000', 'area_direction_layout': ['18.00m²', '1间在租', '1室0厅1卫']}
```

2.2 数据处理

2.2.1 说明

先对数据进行去重，因为爬取还是会不小心爬取广告内容，且有些小区是跨区的。
然后根据字段特性找到面积，朝向，布局位置。
去掉单位。

2.2.2 代码及结果

```
import json
from tqdm import tqdm

# Specify the city
city = "shenzhen"

# Read the data
with open(f"spiderall/3_spider/lianjia_{city}.json", "r") as f:
    data = [json.loads(line) for line in f.readlines()]

# 先去重，所有字段都相同的数据，只保留一条
# 1. 定义一个空列表，用于存放去重后的数据
data_processed = []
```



```

# 2. 遍历 data, 如果 data_processed 中没有该条数据, 则加入 data_processed
for line in tqdm(data, desc="Deduplication", ncols=80):
    if line not in data_processed:
        data_processed.append(line)
# 3. 将 data_processed 赋值给 data
data = data_processed

total_lines = len(data)
processed_lines = 0

for line in tqdm(data, desc="Processing", ncols=80):
    # 1. 处理数据, area_direction_layout 字段, 将其拆分为 area、direction、
    layout 三个字段
    area_direction_layout = line["area_direction_layout"]
    # 三个字段位置不确定
    # area 字段的特点是带有m², direction 字段的特点是带有东西南北, layout 字段
    的特点是带有室厅卫, 所以可以通过这些特点来拆分
    # area 字段
    area = ""
    for i in area_direction_layout:
        if "m²" in i:
            area = i
            break
    # 去掉m², 转换为 float 类型, 20.28-21.14 这种情况, 取中位数
    area = area.replace("m²", "")
    if "-" in area:
        area = float(area.split("-")[0]) + float(area.split("-")[1]) / 2
    else:
        try:
            area = float(area)
        except ValueError:
            area = None
    # direction 字段
    direction = ""
    for i in area_direction_layout:
        if "东" in i or "西" in i or "南" in i or "北" in i:
            direction = i
            break
    # layout 字段
    layout = ""
    for i in area_direction_layout:
        if "室" in i or "厅" in i or "卫" in i:
            layout = i
            break

```

```

# 2. 处理数据, total_price 字段, 将其转换为 int 类型
# "total_price": "5900-6100" 这种情况, 取中位数
total_price = line["total_price"]
if "-" in total_price:
    total_price = (
        int(total_price.split("-")[0]) + int(total_price.split("-")
    )[1]) / 2
)
else:
    total_price = int(total_price)

# 3. 加入 price_per_m2 字段, 保留两位小数
if area != 0 and area is not None:
    price_per_m2 = round(total_price / area, 2)
else:
    price_per_m2 = None

line["price_per_m2"] = price_per_m2

line["area"] = area
line["direction"] = direction
line["layout"] = layout
line["total_price"] = total_price
line.pop("area_direction_layout")

processed_lines += 1

if processed_lines % 100 == 0:
    tqdm.write(f"Processed {processed_lines}/{total_lines} lines")

# 写入数据, lianjia_beijing_processed.json

with open(f"lianjia_{city}_processed.json", "w") as f:
    for line in data:
        f.write(json.dumps(line, ensure_ascii=False))
        f.write("\n")

```

```
spiderall > $ spider > {} ljlajia_beijing.json > ...
1  {"name": "整租·安外东河沿 1室1厅 南", "district": "安定门", "total_price": "6000", "area_direction_layout": ["-", "-", "39.00m\u00b2", "南", "1室1厅1卫"]}
2  {"name": "整租·安德路55号院 2室1厅 北", "district": "安定门", "total_price": "7330", "area_direction_layout": ["|", "南", "48.21m\u00b2", "北", "2室1厅1卫"]}
3  {"name": "整租·安外西河沿路1号院 3室1厅 南", "district": "安定门", "total_price": "17000", "area_direction_layout": ["-", "-", "113.00m\u00b2", "南", "3室1厅1卫"]}
4  {"name": "整租·共享公寓 金台·共享层 套间 1室1厅", "district": null, "total_price": "18000", "area_direction_layout": ["42.86m\u00b2", "南", "1室1厅1卫"]}
5  {"name": "整租·安外西河沿路 1室1厅 南", "district": "安定门", "total_price": "5900", "area_direction_layout": ["-", "-", "43.39m\u00b2", "南", "1室1厅1卫"]}
6  {"name": "整租·城家公寓 马甸康寺大酒店 1个月起租, 可月付 开间", "district": null, "total_price": "5100-5600", "area_direction_layout": ["21.71-23.06m\u00b2", "8间在租", "1室0厅1卫"]}
7  {"name": "整租·六铺炕 3室1厅 南/北", "district": "安定门", "total_price": "13000", "area_direction_layout": ["-", "-", "73.00m\u00b2", "南 北", "3室0厅1卫"]}
8  {"name": "整租·共享层公寓 金台·共享层 开间 开间", "district": null, "total_price": "7500-13000", "area_direction_layout": ["20.98-24.82m\u00b2", "5间在租", "1室0厅1卫"]}
9  {"name": "整租·上龙西里 2室1厅 南", "district": "安定门", "total_price": "7000", "area_direction_layout": ["-", "-", "61.08m\u00b2", "南", "2室1厅1卫"]}
10 {"name": "整租·城家公寓 马甸康寺大酒店 精装直租, 可月付 开间", "district": null, "total_price": "4850-5150", "area_direction_layout": ["22.86-23.86m\u00b2", "2间在租", "1室0厅1卫"]}
11 {"name": "整租·安德路4号院 3室0厅 南/北", "district": "安定门", "total_price": "7200", "area_direction_layout": ["-", "-", "56.65m\u00b2", "南 北", "3室0厅1卫"]}
12 {"name": "整租·城家公寓 马甸康寺大酒店 随时特租直租可做 开间", "district": null, "total_price": "5500-5700", "area_direction_layout": ["22.86m\u00b2", "8间在租", "1室0厅1卫"]}
13 {"name": "整租·西霞房 1室0厅 北", "district": "安定门", "total_price": "5300", "area_direction_layout": ["-", "-", "36.21m\u00b2", "北", "1室0厅1卫"]}
14 {"name": "整租·城家公寓 马甸康寺大酒店 企业特租, 免佣开租 开间", "district": null, "total_price": "5150", "area_direction_layout": ["21.43m\u00b2", "1间在租", "1室0厅1卫"]}
15 {"name": "整租·巷上南园 2室1厅 东/南", "district": "安定门", "total_price": "11600", "area_direction_layout": ["-", "-", "99.27m\u00b2", "东 南", "2室1厅1卫"]}
16 {"name": "整租·安外东河沿 1室1厅 南", "district": "安定门", "total_price": "5800", "area_direction_layout": ["-", "-", "37.64m\u00b2", "南", "1室1厅1卫"]}
17 {"name": "整租·青年湖东里 2室1厅 东/西", "district": "安定门", "total_price": "8000", "area_direction_layout": ["-", "-", "64.20m\u00b2", "东 西", "2室1厅1卫"]}
18 {"name": "整租·城家公寓 马甸康寺大酒店 直租可月付, 可短租 开间", "district": null, "total_price": "5050", "area_direction_layout": ["28.57m\u00b2", "1间在租", "1室0厅1卫"]}
19 {"name": "整租·西霞房 1室1厅 西", "district": "安定门", "total_price": "6700", "area_direction_layout": ["-", "-", "55.95m\u00b2", "西", "1室1厅1卫"]}
20 {"name": "整租·安德路4号院 1室1厅 南", "district": "安定门", "total_price": "5500", "area_direction_layout": ["-", "-", "36.70m\u00b2", "南", "1室1厅1卫"]}
21 {"name": "整租·安德路乙61号院 3室1厅 南/北", "district": "安定门", "total_price": "9800", "area_direction_layout": ["-", "-", "66.67m\u00b2", "南 北", "3室1厅1卫"]}
22 {"name": "整租·共享层公寓 马甸康寺大酒店 近安华桥站, 可月付 开间", "district": null, "total_price": "4050-5150", "area_direction_layout": ["28.57m\u00b2", "2间在租", "1室0厅1卫"]}
23 {"name": "整租·青年湖东里 1室1厅 南", "district": "安定门", "total_price": "6000", "area_direction_layout": ["-", "-", "38.27m\u00b2", "南", "1室1厅1卫"]}
24 {"name": "整租·城家公寓 马甸康寺大酒店 独厨独卫, 精装直租 开间", "district": null, "total_price": "5050-5300", "area_direction_layout": ["22.86m\u00b2", "2间在租", "1室0厅1卫"]}
25 {"name": "整租·黄寺大街28号院 3室1厅 东/北", "district": "安定门", "total_price": "9500", "area_direction_layout": ["-", "-", "81.79m\u00b2", "东 北", "3室1厅1卫"]}
26 {"name": "整租·城家公寓 马甸康寺大酒店 可短租有折扣交通便利 开间", "district": null, "total_price": "5050-5300", "area_direction_layout": ["22.86m\u00b2", "2间在租", "1室0厅1卫"]}
27 {"name": "整租·永康胡同1号院 1室1厅 东/西", "district": "安定门", "total_price": "6300", "area_direction_layout": ["-", "-", "47.00m\u00b2", "东 西", "1室1厅1卫"]}
28 {"name": "整租·城家公寓 马甸康寺大酒店 近地铁精装直租特租 开间", "district": null, "total_price": "4750-6050", "area_direction_layout": ["23.43-30.71m\u00b2", "2间在租", "1室0厅1卫"]}
29 {"name": "整租·花园胡同 3室1厅 东", "district": "安定门", "total_price": "8500", "area_direction_layout": ["-", "-", "59.68m\u00b2", "东", "3室1厅1卫"]}
30 {"name": "整租·城家公寓 马甸康寺大酒店 近地铁可短租有折扣 开间", "district": null, "total_price": "5100-5700", "area_direction_layout": ["23.43-26.43m\u00b2", "4间在租", "1室0厅1卫"]}
31 {"name": "整租·黄寺大街28号院 3室1厅 东/北", "district": "安定门", "total_price": "9500", "area_direction_layout": ["-", "-", "81.79m\u00b2", "东 北", "3室1厅1卫"]}
32 {"name": "整租·城家公寓 马甸康寺大酒店 近地铁一站 酒店式公寓 开间", "district": null, "total_price": "4700", "area_direction_layout": ["25.00m\u00b2", "1间在租", "1室0厅1卫"]}
33 {"name": "整租·安德路47号院 1室1厅 南", "district": "安定门", "total_price": "6500", "area_direction_layout": ["-", "-", "49.21m\u00b2", "南", "1室1厅1卫"]}
34 {"name": "整租·安外东河沿 1室1厅 南", "district": "安定门", "total_price": "3000", "area_direction_layout": ["18.00m\u00b2", "1间在租", "1室0厅1卫"]}
35 {"name": "整租·西霞房 2室1厅 西", "district": "安定门", "total_price": "9000", "area_direction_layout": ["-", "-", "85.00m\u00b2", "西", "2室1厅1卫"]}
36 {"name": "整租·城家公寓 马甸康寺大酒店 独厨独卫精装, 直租 开间", "district": null, "total_price": "5150", "area_direction_layout": ["23.43m\u00b2", "1间在租", "1室0厅1卫"]}
37 {"name": "整租·永康胡同1号院 1室1厅 东/西", "district": "安定门", "total_price": "6300", "area_direction_layout": ["-", "-", "47.00m\u00b2", "东 西", "1室1厅1卫"]}
38 {"name": "整租·城家公寓 马甸康寺大酒店 近地铁一站 loft公寓精装修介绍 1室1厅", "district": null, "total_price": "4000", "area_direction_layout": ["25.00m\u00b2", "1间在租", "1室0厅1卫"]}
39 {"name": "整租·花园胡同 3室1厅 东", "district": "安定门", "total_price": "8500", "area_direction_layout": ["-", "-", "59.68m\u00b2", "东", "3室1厅1卫"]}
40 {"name": "整租·城家公寓 马甸康寺大酒店 近地铁 随时特租 开间", "district": null, "total_price": "5100-5700", "area_direction_layout": ["23.43-26.43m\u00b2", "4间在租", "1室0厅1卫"]}
41 {"name": "整租·安德路47号院 1室1厅 南", "district": "安定门", "total_price": "6500", "area_direction_layout": ["-", "-", "49.21m\u00b2", "南", "1室1厅1卫"]}
42 {"name": "整租·城家公寓 马甸康寺大酒店 可短租有折扣出行便利 开间", "district": null, "total_price": "5500-5600", "area_direction_layout": ["26.43m\u00b2", "4间在租", "1室0厅1卫"]}
43 {"name": "整租·郎家胡同 1室1厅 南", "district": "安定门", "total_price": "4100", "area_direction_layout": ["-", "-", "16.30m\u00b2", "南", "1室1厅1卫"]}
44 {"name": "整租·城家公寓 马甸康寺大酒店 近地铁精装可短租 开间", "district": null, "total_price": "5400-5600", "area_direction_layout": ["22.86m\u00b2", "6间在租", "1室0厅1卫"]}
45 {"name": "整租·安外东河沿 1室1厅 南/北", "district": "安定门", "total_price": "8000", "area_direction_layout": ["-", "-", "54.46m\u00b2", "南 北", "1室1厅1卫"]}
46 {"name": "整租·砂锅胡同 1室0厅 西", "district": "安定门", "total_price": "2800", "area_direction_layout": ["-", "-", "13.25m\u00b2", "西", "1室0厅1卫"]}
47 {"name": "整租·安外西河沿路2号院 车库", "district": "安定门", "total_price": "14900", "area_direction_layout": ["-", "-", "83.85m\u00b2", "南", "1室0厅1卫"]}
48 {"name": "整租·城家公寓 马甸康寺大酒店 随时特租可短租 开间", "district": null, "total_price": "5500-5600", "area_direction_layout": ["22.57-23.06m\u00b2", "3间在租", "1室0厅1卫"]}
49 {"name": "整租·安外东河沿 1室1厅 南", "district": "安定门", "total_price": "6100", "area_direction_layout": ["-", "-", "41.52m\u00b2", "南", "1室1厅1卫"]}
50 {"name": "整租·安德路47号院 1室1厅 东/西", "district": "安定门", "total_price": "6900", "area_direction_layout": ["-", "-", "51.01m\u00b2", "东 西", "2室1厅1卫"]}
51 {"name": "整租·安外大街3号院 2室1厅 南/北", "district": "安定门", "total_price": "6500", "area_direction_layout": ["-", "-", "55.20m\u00b2", "南 北", "2室1厅1卫"]}
52 {"name": "整租·六铺炕 1室1厅 南", "district": "安定门", "total_price": "5000", "area_direction_layout": ["|", "南", "36.53m\u00b2", "南", "1室1厅1卫"]}
53 {"name": "整租·青年湖东里 2室1厅 西/北", "district": "安定门", "total_price": "8000", "area_direction_layout": ["-", "-", "61.15m\u00b2", "西 北", "2室1厅1卫"]}
54 {"name": "整租·安德路乙61号院 2室1厅 东", "district": "安定门", "total_price": "7530", "area_direction_layout": ["|", "南", "54.69m\u00b2", "南", "2室1厅1卫"]}
55 {"name": "整租·安外西河沿路1号院 1室1厅 南", "district": "安定门", "total_price": "11500", "area_direction_layout": ["-", "-", "65.00m\u00b2", "南", "1室1厅1卫"]}
```

↑原数据

```
{
  "ljlajia_beijing_processed.json": [
    {
      "name": "整租·安外东河沿 1室1厅 南", "district": "安定门", "total_price": 6000, "price_per_m2": 153.85, "area": 39.0, "direction": "南", "layout": "1室1厅1卫"},
      {
        "name": "整租·安德路55号院 2室1厅 北", "district": "安定门", "total_price": 7330, "price_per_m2": 152.04, "area": 48.21, "direction": "北", "layout": "2室1厅1卫"},
        {
          "name": "整租·安外西河沿路1号院 3室1厅 南", "district": "安定门", "total_price": 17000, "price_per_m2": 150.44, "area": 113.0, "direction": "南", "layout": "3室1厅1卫"},
          {
            "name": "整租·共享公寓 金台·共享层 套间 1室1厅", "district": null, "total_price": 18000, "price_per_m2": 419.97, "area": 42.86, "direction": "", "layout": "1室1厅1卫"},
            {
              "name": "整租·安外西河沿 1室1厅 南", "district": "安定门", "total_price": 5900, "price_per_m2": 135.98, "area": 43.39, "direction": "南", "layout": "1室1厅1卫"},
              {
                "name": "整租·城家公寓 马甸康寺大酒店 1个月起租, 可月付 开间", "district": null, "total_price": 7900.0, "price_per_m2": 234.84, "area": 33.64, "direction": "", "", "layout": "1室0厅1卫"},
                {
                  "name": "整租·六铺炕 3室1厅 南/北", "district": "安定门", "total_price": 13000, "price_per_m2": 178.08, "area": 73.0, "direction": "南 北", "layout": "3室0厅1卫"},
                  {
                    "name": "整租·安外东河沿 1室1厅 南", "district": "安定门", "total_price": 5800, "price_per_m2": 154.09, "area": 37.64, "direction": "南", "layout": "1室1厅1卫"},
                    {
                      "name": "整租·上龙西里 2室1厅 南", "district": "安定门", "total_price": 7000, "price_per_m2": 114.6, "area": 61.08, "direction": "南", "layout": "2室1厅1卫"},
                      {
                        "name": "整租·城家公寓 马甸康寺大酒店 精装直租, 可月付 开间", "district": null, "total_price": 7425.0, "price_per_m2": 213.42, "area": 34.79, "direction": "", "layout": "1室0厅1卫"},
                        {
                          "name": "整租·安德路47号院 3室0厅 南/北", "district": "安定门", "total_price": 7200, "price_per_m2": 127.1, "area": 56.65, "direction": "南 北", "layout": "3室0厅1卫"},
                          {
                            "name": "整租·城家公寓 马甸康寺大酒店 随时特租直租可做 开间", "district": null, "total_price": 8350.0, "price_per_m2": 365.27, "area": 22.86, "direction": "", "layout": "1室0厅1卫"},
                            {
                              "name": "整租·西霞房 1室0厅 北", "district": "安定门", "total_price": 5300, "price_per_m2": 145.97, "area": 36.31, "direction": "北", "layout": "1室0厅1卫"},
                              {
                                "name": "整租·城家公寓 马甸康寺大酒店 企业特租, 免佣开租 开间", "district": null, "total_price": 5150, "price_per_m2": 240.32, "area": 21.43, "direction": "", "", "layout": "1室0厅1卫"},
                                {
                                  "name": "整租·巷上南园 2室1厅 东/南", "district": "安定门", "total_price": 11600, "price_per_m2": 116.85, "area": 99.27, "direction": "东 南", "layout": "2室1厅1卫"},
                                  {
                                    "name": "整租·安外东河沿 1室1厅 南", "district": "安定门", "total_price": 5800, "price_per_m2": 154.09, "area": 37.64, "direction": "南", "layout": "1室1厅1卫"},
                                    {
                                      "name": "整租·青年湖东里 1室1厅 南", "district": "安定门", "total_price": 6000, "price_per_m2": 124.61, "area": 64.2, "direction": "东 西", "layout": "2室1厅1卫"},
                                      {
                                        "name": "整租·城家公寓 马甸康寺大酒店 直租可月付, 可短租 开间", "district": null, "total_price": 5050, "price_per_m2": 176.76, "area": 28.57, "direction": "", "layout": "1室0厅1卫"},
                                        {
                                          "name": "整租·西霞房 1室1厅 西", "district": "安定门", "total_price": 6700, "price_per_m2": 119.75, "area": 55.95, "direction": "西", "layout": "1室1厅1卫"},
                                          {
                                            "name": "整租·安德路47号院 1室1厅 南", "district": "安定门", "total_price": 5500, "price_per_m2": 149.86, "area": 36.7, "direction": "南", "layout": "1室1厅1卫"},
                                            {
                                              "name": "整租·安德路乙61号院 3室1厅 南/北", "district": "安定门", "total_price": 9800, "price_per_m2": 146.99, "area": 66.67, "direction": "南 北", "layout": "3室1厅1卫"},
                                              {
                                                "name": "整租·城家公寓 马甸康寺大酒店 近安华桥站, 可月付 开间", "district": null, "total_price": 7525.0, "price_per_m2": 263.39, "area": 28.57, "direction": "", "layout": "1室0厅1卫"},
                                                {
                                                  "name": "整租·青年湖东里 1室1厅 南", "district": "安定门", "total_price": 6000, "price_per_m2": 156.78, "area": 38.27, "direction": "南", "layout": "1室1厅1卫"},
                                                  {
                                                    "name": "整租·城家公寓 马甸康寺大酒店 独厨独卫, 精装直租 开间", "district": null, "total_price": 7700.0, "price_per_m2": 336.83, "area": 22.86, "direction": "", "layout": "1室0厅1卫"},
                                                    {
                                                      "name": "整租·城家公寓 马甸康寺大酒店 近地铁 随时特租 开间", "district": "安定门", "total_price": 7000, "price_per_m2": 114.6, "area": 61.08, "direction": "南", "layout": "2室1厅1卫"},
                                                      {
                                                        "name": "整租·城家公寓 马甸康寺大酒店 可短租有折扣交通便利 开间", "district": null, "total_price": 7700.0, "price_per_m2": 336.83, "area": 22.86, "direction": "", "layout": "1室0厅1卫"},
                                                        {
                                                          "name": "整租·永康胡同1号院 1室1厅 东/西", "district": "安定门", "total_price": 6300, "price_per_m2": 134.04, "area": 47.0, "direction": "东 西", "layout": "1室1厅1卫"},
                                                          {
                                                            "name": "整租·城家公寓 马甸康寺大酒店 近地铁精装直租特租 开间", "district": null, "total_price": 7775.0, "price_per_m2": 200.46, "area": 38.785, "direction": "", "layout": "1室0厅1卫"},
                                                            {
                                                              "name": "整租·花园胡同 3室1厅 东", "district": "安定门", "total_price": 8500, "price_per_m2": 142.43, "area": 59.68, "direction": "东", "layout": "3室1厅1卫"},
                                                              {
                                                                "name": "整租·城家公寓 马甸康寺大酒店 近地铁可短租有折扣 开间", "district": null, "total_price": 7950.0, "price_per_m2": 216.95, "area": 36.649999999999996, "direction": "", "layout": "1室0厅1卫"},
                                                                {
                                                                  "name": "整租·城家公寓 马甸康寺大酒店 近地铁一站 酒店式公寓 开间", "district": null, "total_price": 4700, "price_per_m2": 188.0, "area": 25.0, "direction": "", "layout": "1室0厅1卫"},
                                                                  {
                                                                    "name": "整租·安德路47号院 1室1厅 南", "district": "安定门", "total_price": 6500, "price_per_m2": 132.09, "area": 49.21, "direction": "南", "layout": "1室1厅1卫"},
                                                                    {
                                                                      "name": "整租·城家公寓 马甸康寺大酒店 近地铁 随时特租 开间", "district": null, "total_price": 5100-5700, "price_per_m2": 166.67, "area": 18.0, "direction": "", "layout": "1室0厅1卫"},
                                                                      {
                                                                        "name": "整租·西霞房 2室1厅 西", "district": "安定门", "total_price": 9000, "price_per_m2": 105.88, "area": 85.0, "direction": "西", "layout": "2室1厅1卫"},
                                                                        {
                                                                          "name": "整租·城家公寓 马甸康寺大酒店 独厨独卫精装, 直租 开间", "district": null, "total_price": 5150, "price_per_m2": 219.8, "area": 23.43, "direction": "", "layout": "1室0厅1卫"},
                                                                          {
                                                                            "name": "整租·城家公寓 马甸康寺大酒店 近地铁精装直租, 限时特租 开间", "district": null, "total_price": 5400, "price_per_m2": 160.0, "area": 25.0, "direction": "", "layout": "1室0厅1卫"},
                                                                            {
                                                                              "name": "整租·城家公寓 马甸康寺大酒店 近地铁精装可短租 开间", "district": null, "total_price": 8150.0, "price_per_m2": 227.75, "area": 35.785, "direction": "", "layout": "1室0厅1卫"},
                                                                              {
                                                                                "name": "整租·安德路47号院 1室1厅 南/北", "district": "安定门", "total_price": 8500, "price_per_m2": 119.72, "area": 71.0, "direction": "南 北", "layout": "3室1厅1卫"},
                                                                                {
                                                                                  "name": "整租·城家公寓 马甸康寺大酒店 可短租有折扣出行便利 开间", "district": null, "total_price": 8300.0, "price_per_m2": 314.04, "area": 26.43, "direction": "", "layout": "1室0厅1卫"},
                                                                                  {
                                                                                    "name": "整租·郎家胡同 1室1厅 南", "district": "安定门", "total_price": 4100, "price_per_m2": 251.53, "area": 16.3, "direction": "南", "layout": "1室1厅1卫"},
                                                                                    {
                                                                                      "name": "整租·城家公寓 马甸康寺大酒店 近地铁精装可短租 开间", "district": null, "total_price": 8200.0, "price_per_m2": 358.71, "area": 22.86, "direction": "", "layout": "1室0厅1卫"},
                                                                                      {
                                                                                        "name": "整租·安外东河沿 1室1厅 南/北", "district": "安定门", "total_price": 8000, "price_per_m2": 146.9, "area": 54.46, "direction": "南 北", "layout": "1室1厅1卫"},
                                                                                        {
                                                                                          "name": "整租·砂锅胡同 1室0厅 西", "district": "安定门", "total_price": 2800, "price_per_m2": 211.32, "area": 13.25, "direction": "西", "layout": "1室0厅1卫"},
                                                                                          {
                                                                                            "name": "整租·安外西河沿路2号院 车库", "district": "安定门", "total_price": 14900, "price_per_m2": 177.7, "area": 83.85, "direction": "", "layout": ""},
                                                                                            {
                                                                                              "name": "整租·城家公寓 马甸康寺大酒店 随时特租可短租 开间", "district": null, "total_price": 8300.0, "price_per_m2": 243.62, "area": 34.07, "direction": "", "", "layout": "1室0厅1卫"},
                                                                                              {
                                                                                                "name": "整租·安外东河沿 1室1厅 南", "district": "安定门", "total_price": 6100, "price_per_m2": 146.92, "area": 41.52, "direction": "南", "layout": "1室1厅1卫"},
                                                                                                {
                                                                                                  "name": "整租·安德路47号院 1室1厅 东/西", "district": "安定门", "total_price": 6900, "price_per_m2": 135.27, "area": 51.01, "direction": "东 西", "layout": "2室1厅1卫"},
                                                                                                  {
                                                                                                    "name": "整租·安外大街3号院 2室1厅 南/北", "district": "安定门", "total_price": 6500, "price_per_m2": 117.75, "area": 55.2, "direction": "南 北", "layout": "2室1厅1卫"},
                                                                                                    {
                                                                                                      "name": "整租·六铺炕 1室1厅 南", "district": "安定门", "total_price": 5000, "price_per_m2": 136.87, "area": 36.53, "direction": "南", "layout": "1室1厅1卫"},
                                                                                                      {
                                                                                                        "name": "整租·青年湖东里 2室1厅 西/北", "district": "安定门", "total_price": 8000, "price_per_m2": 137.69, "area": 61.15, "direction": "西 北", "layout": "2室1厅1卫"},
                                                                                                        {
                                                                                                          "name": "整租·安德路乙61号院 2室1厅 东", "district": "安定门", "total_price": 7530, "price_per_m2": 137.69, "area": 54.69, "direction": "东", "layout": "2室1厅1卫"},
                                                                                                          {
                                                                                                            "name": "整租·安外西河沿路1号院 1室1厅 南", "district": "安定门", "total_price": 11500, "price_per_m2": 176.92, "area": 65.0, "direction": "南", "layout": "1室1厅1卫"},
                                                                                                            {
                                                                                                              "name": "整租·青年湖东里 2室1厅 南", "district": "安定门", "total_price": 7360, "price_per_m2": 132.8, "area": 55.42, "direction": "南", "layout": "2室1厅1卫"},
                                                                                                              {
                                                                                                                "name": "整租·上龙西里 3室1厅 南/北", "district": "安定门", "total_price": 6900, "price_per_m2": 116.95, "area": 59.0, "direction": "南 北", "layout": "3室1厅1卫"},
                                                                                                                {
                                                                                                                  "name": "整租·安外东河沿 1室1厅 南", "district": "安定门", "total_price": 6200, "price_per_m2": 180.40, "area": 37.62, "direction": "南", "layout": "1室1厅1卫"}

```

↑处理后的数据

2.3 数据可视化

2.3.1 说明

1. 比较 5 个城市的总体房租情况，包含租金的均价、最高价、最低价、中位数等信息，单位面积租金（元/平米）的均价、最高价、最低价、中位数等信息。
2. 比较 5 个城市一居、二居、三居的情况，包含均价、最高价、最低价、中位数等信息，采用合适的图或表形式进行展示。
3. 计算和分析每个城市不同板块的均价情况，并采用合适的图或表形式进行展示。
4. 比较各个城市不同朝向的单位面积租金分布情况，采用合适的图或表形式进行展示。
5. 查询各个城市的平均工资，分析并展示其和单位面积租金分布的关系。比较一下在哪个城市租房的负担最重？
6. 与 2022 年的租房数据进行对比（只比较北上广深 4 个城市，原始数据会给出），总结你观察到的变化情况，并用图、表、文字等支撑你得到的结论。
7. 为了更好分析板块均价，我采用热力图形式可视化。

2.3.2 代码

比较 5 个城市的总体房租情况，包含租金的均价、最高价、最低价、中位数等信息，单位面积租金（元/平米）的均价、最高价、最低价、中位数等信息。采用合适的图或表形式进行展示。

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

matplotlib.use("TKAgg")
import seaborn as sns

cities = ["beijing", "shanghai", "shenzhen", "guangzhou", "suzhou"]

# 读取数据,lianjia_{city}_processed.json
data = pd.DataFrame()
for city in cities:
    file_path = f"lianjia_{city}_processed.json"
    city_data = pd.read_json(file_path, lines=True)
    city_data["city"] = city # Add city column
    data = pd.concat([data, city_data])
```

```
# {"name": "整租·方丹苑 2室1厅 东北", "district": "潘家园",  
"total_price": 8200, "price_per_m2": 92.13, "area": 89.0, "direction":  
"东北", "layout": "2室1厅1卫"}
```

```
# 存在没有总价的情况，将其删除
```

```
data["total_price"] = pd.to_numeric(data["total_price"],  
errors="coerce")
```

```
# # 找出包含字符串值的行
```

```
# string_values = data[pd.isna(data['total_price'])]
```

```
# print(string_values)
```

```
# 租金的均价，total_price 字段
```

```
avg_total_price =  
data.groupby("city")["total_price"].mean().reset_index()  
avg_total_price.columns = ["city", "avg_total_price"]
```

```
# 租金的最高价，total_price 字段
```

```
max_total_price =  
data.groupby("city")["total_price"].max().reset_index()  
max_total_price.columns = ["city", "max_total_price"]
```

```
# 租金的最低价，total_price 字段
```

```
min_total_price =  
data.groupby("city")["total_price"].min().reset_index()  
min_total_price.columns = ["city", "min_total_price"]
```

```
# 租金的中位数，total_price 字段
```

```
median_total_price =  
data.groupby("city")["total_price"].median().reset_index()  
median_total_price.columns = ["city", "median_total_price"]
```

```
# 单位面积租金的均价，price_per_m2 字段
```

```
avg_price_per_m2 =  
data.groupby("city")["price_per_m2"].mean().reset_index()  
avg_price_per_m2.columns = ["city", "avg_price_per_m2"]
```

```
# 单位面积租金的最高价，price_per_m2 字段
```

```
max_price_per_m2 =  
data.groupby("city")["price_per_m2"].max().reset_index()  
max_price_per_m2.columns = ["city", "max_price_per_m2"]
```

```
# 单位面积租金的最低价，price_per_m2 字段
```

```
min_price_per_m2 =  
data.groupby("city")["price_per_m2"].min().reset_index()
```

```

min_price_per_m2.columns = ["city", "min_price_per_m2"]

# 单位面积租金的中位数, price_per_m2 字段
median_price_per_m2 =
data.groupby("city")["price_per_m2"].median().reset_index()
median_price_per_m2.columns = ["city", "median_price_per_m2"]

# 采用合适的图或表形式进行展示
# total_price
# 柱状图
fig, axes = plt.subplots(2, 2, figsize=(12, 8))
sns.barplot(x="city", y="avg_total_price", data=avg_total_price,
ax=axes[0][0])
sns.barplot(x="city", y="max_total_price", data=max_total_price,
ax=axes[0][1])
sns.barplot(x="city", y="min_total_price", data=min_total_price,
ax=axes[1][0])
sns.barplot(x="city", y="median_total_price", data=median_total_price,
ax=axes[1][1])

plt.show()

# price_per_m2
# 柱状图
fig, axes = plt.subplots(2, 2, figsize=(12, 8))
sns.barplot(x="city", y="avg_price_per_m2", data=avg_price_per_m2,
ax=axes[0][0])
sns.barplot(x="city", y="max_price_per_m2", data=max_price_per_m2,
ax=axes[0][1])
sns.barplot(x="city", y="min_price_per_m2", data=min_price_per_m2,
ax=axes[1][0])
sns.barplot(x="city", y="median_price_per_m2",
data=median_price_per_m2, ax=axes[1][1])

plt.show()

```

比较 5 个城市一居、二居、三居的情况, 包含均价、最高价、最低价、中位数等信息, 采用合适的图或表形式进行展示。

```

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

```

```

matplotlib.use("TKAgg")
import seaborn as sns

plt.rcParams["font.sans-serif"] = ["SimHei"]

cities = ["beijing", "shanghai", "shenzhen", "guangzhou", "suzhou"]

# 读取数据, lianjia_{city}_processed.json
data = pd.DataFrame()
for city in cities:
    file_path = f"lianjia_{city}_processed.json"
    city_data = pd.read_json(file_path, lines=True)
    city_data["city"] = city # Add city column
    data = pd.concat([data, city_data])

# 因为根据房型来分析, 所以需要先分为整租、合租、独栋三种租赁类型
data["rent_type"] = ""
# 整租
data.loc[data["name"].str.contains("整租"), "rent_type"] = "整租"
# 合租
data.loc[data["name"].str.contains("合租"), "rent_type"] = "合租"
# 独栋
data.loc[data["name"].str.contains("独栋"), "rent_type"] = "独栋"
# 都不是
data.loc[data["rent_type"] == "", "rent_type"] = "其他"

# 比较同一租赁方式, 同一房型在不同城市的数量、均价、最高价、最低价、中位数等信息, 采用合适的图或表形式进行展示

# 不同房型还是需要聚合一下, 比如 1 室 1 厅、1 室 0 厅、1 室 2 厅都是 1 室, 所以需要把这些房型都归为 1 室

nan_rows = data[data["layout"].isna()]
print(nan_rows)

# 1 室或 1 房间
data.loc[data["layout"].str.contains("1 室", na=False), "layout"] = "1 居室"
data.loc[data["layout"].str.contains("1 房间", na=False), "layout"] = "1 居室"
# 2 室或 2 房间
data.loc[data["layout"].str.contains("2 室", na=False), "layout"] = "2 居室"

```



```

data.loc[data["layout"].str.contains("2 房间", na=False), "layout"] = "2
居室"
# 3 室或 3 房间
data.loc[data["layout"].str.contains("3 室", na=False), "layout"] = "3 居
室"
data.loc[data["layout"].str.contains("3 房间", na=False), "layout"] = "3
居室"

# Group the data and calculate the statistics
grouped_data = data.groupby(["rent_type", "city",
"layout"])[["total_price"]].agg(
    ["count", "mean", "max", "min", "median"]
)

# Reset the index to make 'rent_type', 'city', and 'layout' columns
again
grouped_data = grouped_data.reset_index()

# Filter the layout to only include 1 居室, 2 居室, and 3 居室
filtered_data = grouped_data[grouped_data["layout"].isin(["1 居室", "2 居
室", "3 居室"])]

# Print the table
# print(filtered_data)

def plot_rent_type(filtered_data, rent_type):
    fig, axes = plt.subplots(2, 2, figsize=(12, 8))
    sns.barplot(
        x="city",
        y="mean",
        hue="layout",
        data=filtered_data[filtered_data["rent_type"] == rent_type],
        ax=axes[0][0],
    )
    axes[0][0].set_title(f"{rent_type} - 均价")
    sns.barplot(
        x="city",
        y="max",
        hue="layout",
        data=filtered_data[filtered_data["rent_type"] == rent_type],
        ax=axes[0][1],
    )
    axes[0][1].set_title(f"{rent_type} - 最高价")

```



```
sns.barplot(
    x="city",
    y="min",
    hue="layout",
    data=filtered_data[filtered_data["rent_type"] == rent_type],
    ax=axes[1][0],
)
axes[1][0].set_title(f"{rent_type} - 最低价")
sns.barplot(
    x="city",
    y="median",
    hue="layout",
    data=filtered_data[filtered_data["rent_type"] == rent_type],
    ax=axes[1][1],
)
axes[1][1].set_title(f"{rent_type} - 中位数")
plt.show()
```

```
plot_rent_type(filtered_data, "整租")
# plot_rent_type(filtered_data, "合租")
# plot_rent_type(filtered_data, "独栋")
```

```

display3.py > plot_mean_price
1  # 计算和分析每个城市不同板块的均价情况，并采用合适的图或表形式进行展示。
2
3  import pandas as pd
4  import matplotlib.pyplot as plt
5  import matplotlib
6  import seaborn as sns
7
8  matplotlib.use("TKAgg")
9  plt.rcParams["font.sans-serif"] = ["SimHei"]
10
11
12  def plot_mean_price(city):
13      # Read data from the processed file
14      df = pd.read_json(f"./lianjia_{city}_processed.json", lines=True)
15
16      # Group by district and calculate the mean price and count of houses
17      groupby_district = df.groupby("district")
18      mean_price = groupby_district["price_per_m2"].mean().sort_values(ascending=False)
19      house_count = groupby_district.size().sort_values(ascending=False)
20
21      # Plot the mean price using seaborn
22      fig, ax1 = plt.subplots(figsize=(20, 10))
23      ax1.bar(mean_price.index, mean_price.values, color="blue")
24      ax1.set_xticklabels(mean_price.index, rotation=90, fontsize=4.34)
25      ax1.set_xlabel("板块")
26      ax1.set_ylabel("均价 (元/㎡每月) ")
27      ax1.set_title(f"{city}各个板块的均价和房子数量")
28
29      # Create a secondary y-axis for house count
30      ax2 = ax1.twinx()
31      ax2.bar(house_count.index, house_count.values, alpha=0.5, color="red")
32      ax2.set_ylabel("房子数量")
33
34      plt.show()
35
36      # Get the top 10 and last 10 districts by mean price
37      top_10_districts = mean_price.head(10)
38      last_10_districts = mean_price.tail(10)
39
40      # Plot the top 10 districts
41      plt.figure(figsize=(10, 5))
42      sns.barplot(x=top_10_districts.index, y=top_10_districts.values, color="blue")
43      plt.xticks(rotation=90)
44      plt.xlabel("板块")
45      plt.ylabel("均价 (元/㎡每月) ")
46      plt.title("前十个板块的均价")
47      plt.show()
48
49      # Plot the last 10 districts
50      plt.figure(figsize=(10, 5))
51      sns.barplot(x=last_10_districts.index, y=last_10_districts.values, color="blue")
52      plt.xticks(rotation=90)
53      plt.xlabel("板块")
54      plt.ylabel("均价 (元/㎡每月) ")
55      plt.title("后十个板块的均价")
56      plt.show()
57
58
59  # Call the function for a specific city
60  plot_mean_price("suzhou")
61

```

热力图展示


```

display5.py > ...
1 # 与2022年的租房数据进行对比（只比较北上广深4个城市，原始数据会给出），总结你观察到的变化情况，并用图、表、文字等支撑你得到的结论。
2
3 # 2022年的租房数据 RawData2022/{city}HouseInfo.json
4 # 2023年的租房数据,lianjia_{city}_processed.json
5
6 import pandas as pd
7 import matplotlib.pyplot as plt
8 import matplotlib
9 import seaborn as sns
10
11 matplotlib.use("TKAgg")
12
13 matplotlib.rcParams["font.family"] = "SimHei"
14
15 cities = ["beijing", "shanghai", "guangzhou", "shenzhen"]
16
17 data_2022 = {}
18 data_2023 = {}
19
20 for city in cities:
21     data_2022[city] = pd.read_json(
22         f"RawData2022/{city.capitalize()}HouseInfo.json", lines=True
23     )
24     data_2023[city] = pd.read_json(f"lianjia_{city}_processed.json", lines=True)
25
26 avg_prices_2022 = {city: df["total_price"].mean() for city, df in data_2022.items()}
27 avg_prices_2023 = {
28     city: df[df["name"].str.startswith("整租")]["total_price"].mean()
29     for city, df in data_2023.items()
30 }
31
32 df_2022 = pd.DataFrame.from_dict(
33     avg_prices_2022, orient="index", columns=["Avg Price 2022"]
34 )
35 df_2023 = pd.DataFrame.from_dict(
36     avg_prices_2023, orient="index", columns=["Avg Price 2023"]
37 )
38
39 df = pd.concat([df_2022, df_2023], axis=1)
40
41 df.plot(kind="bar", figsize=(10, 6))
42 plt.title("Average Rental Prices in 2022 vs 2023")
43 plt.ylabel("Average Price")
44 plt.show()
45
46

```

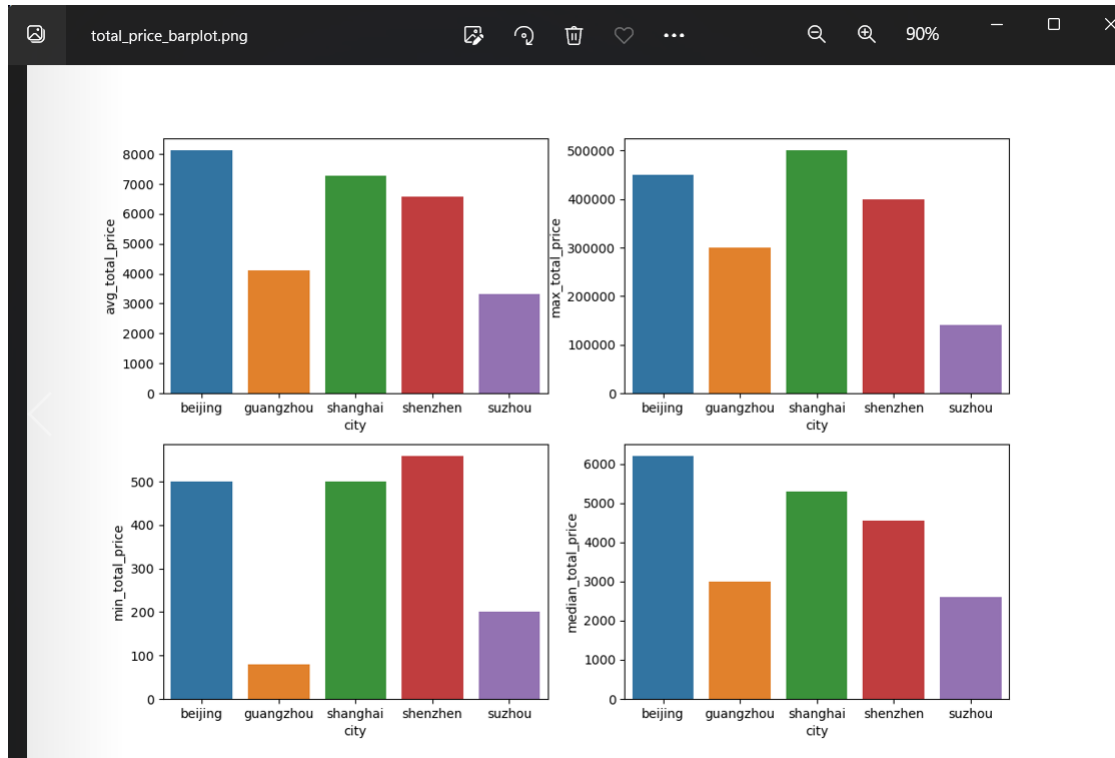
```

display6.py > ...
3 # 北京 13251
4 # 上海 13433
5 # 深圳 12668
6 # 广州 10913
7 # 苏州 11037
8
9
10 import pandas as pd
11 import matplotlib.pyplot as plt
12 import matplotlib
13
14 matplotlib.use("TKAgg")
15 import seaborn as sns
16
17 cities = ["beijing", "shanghai", "shenzhen", "guangzhou", "suzhou"]
18
19 # 读取数据,lianjia_{city}_processed.json
20 data = pd.DataFrame()
21 for city in cities:
22     file_path = f"lianjia_{city}_processed.json"
23     city_data = pd.read_json(file_path, lines=True)
24     city_data["city"] = city # Add city column
25     data = pd.concat([data, city_data])
26
27 # 存在没有总价的情况，将其删除
28 data["total_price"] = pd.to_numeric(data["total_price"], errors="coerce")
29
30 # 单位面积租金的均价，price_per_m2字段
31 avg_price_per_m2 = data.groupby("city")["price_per_m2"].mean().reset_index()
32 avg_price_per_m2.columns = ["city", "avg_price_per_m2"]
33
34 # 月工资数据
35 monthly_salaries = [13251, 13433, 12668, 10913, 11037]
36 avg_price_per_m2["monthly_salary"] = monthly_salaries
37
38 # 采用合适的图或表形式进行展示
39 # price_per_m2
40 # 柱状图
41 fig, axes = plt.subplots(1, 2, figsize=(12, 4))
42
43 # 绘制单位面积租金的均价柱状图
44 sns.barplot(x="city", y="avg_price_per_m2", data=avg_price_per_m2, ax=axes[0])
45 axes[0].set_title("Average Price per m2")
46
47 # 绘制月工资的柱状图
48 sns.barplot(x="city", y="monthly_salary", data=avg_price_per_m2, ax=axes[1])
49 axes[1].set_title("Monthly Salary")
50
51 # 计算单位面积租金和月工资的比值
52 avg_price_per_m2["price_salary_ratio"] = (
53     avg_price_per_m2["avg_price_per_m2"] / avg_price_per_m2["monthly_salary"]
54 )
55
56 # 创建比值表
57 ratio_table = avg_price_per_m2[["city", "price_salary_ratio"]]
58
59 # 打印比值表
60 print(ratio_table)
61
62 plt.show()
63

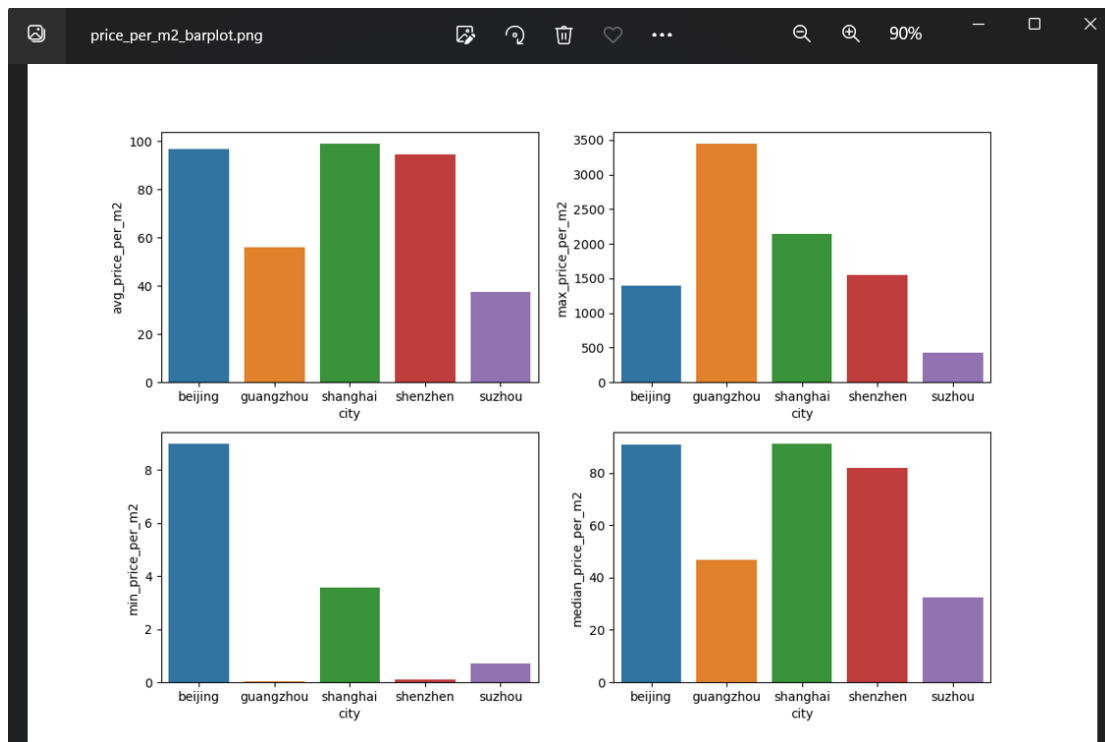
```

3 可视化结果及分析

3.1 总价和平米均价

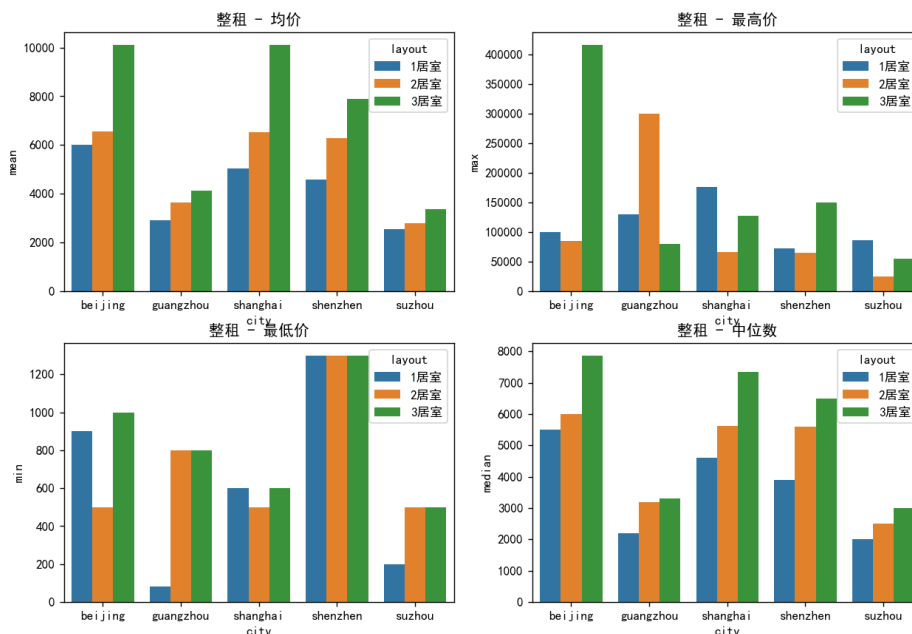


总价显示北京平均总价最高，中位数也最高。广州和苏州租房价格较北京上海深圳低。



每平米单价上来看, 上海总价和中位数最高, 大概是因为上海租房平均面积比北京小。同样, 单价上, 广州和苏州较北京上海深圳低。

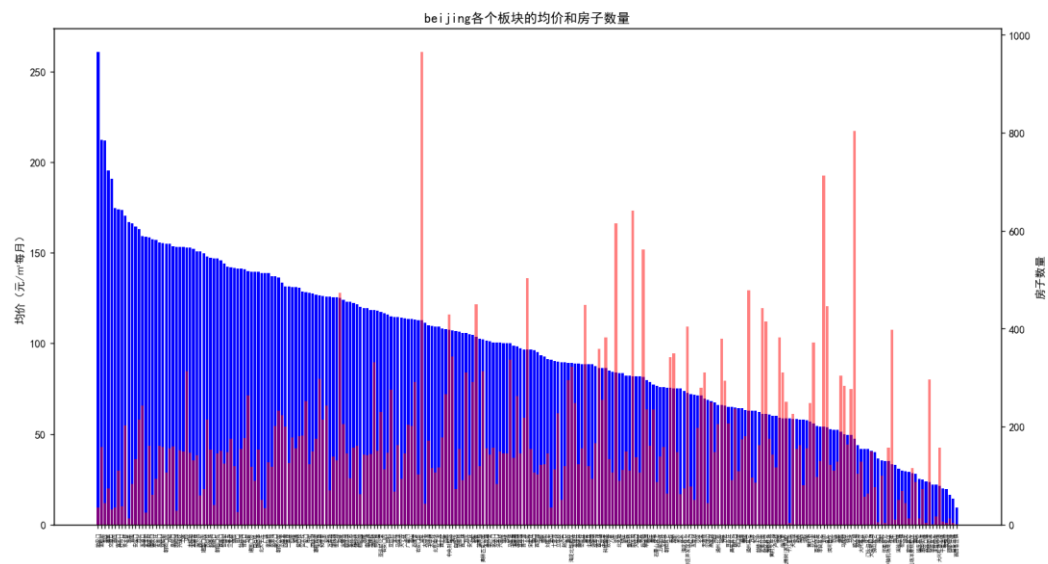
3.2 不同布局总价



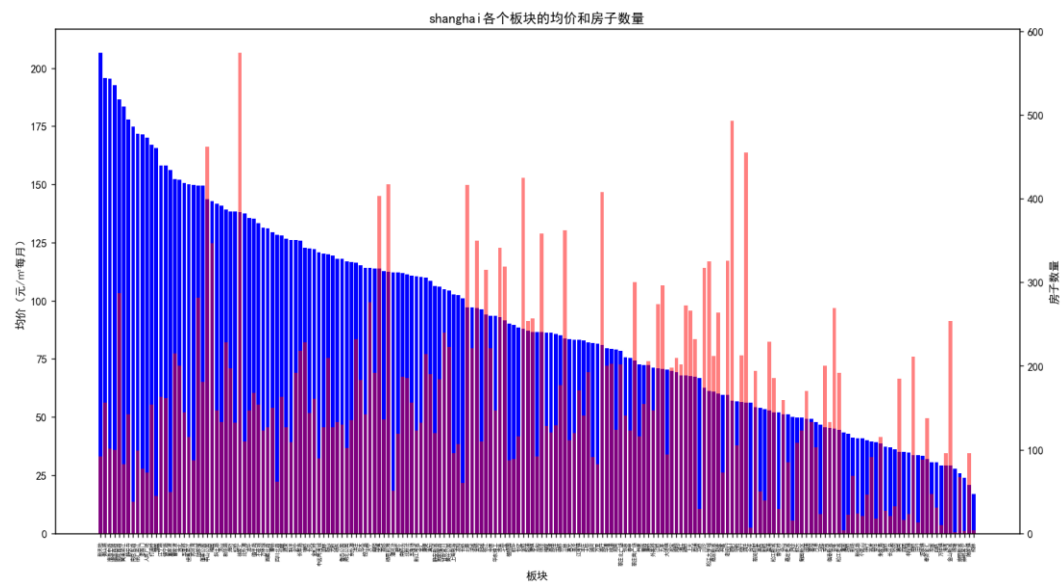
整租租房图表中可以看到 1 居室总价均价上, 北京最贵, 2 居室和 3 居室, 北京上海最贵。总价中位数上, 北京最贵。房价上, 北京>上海>深圳>广州>苏州。

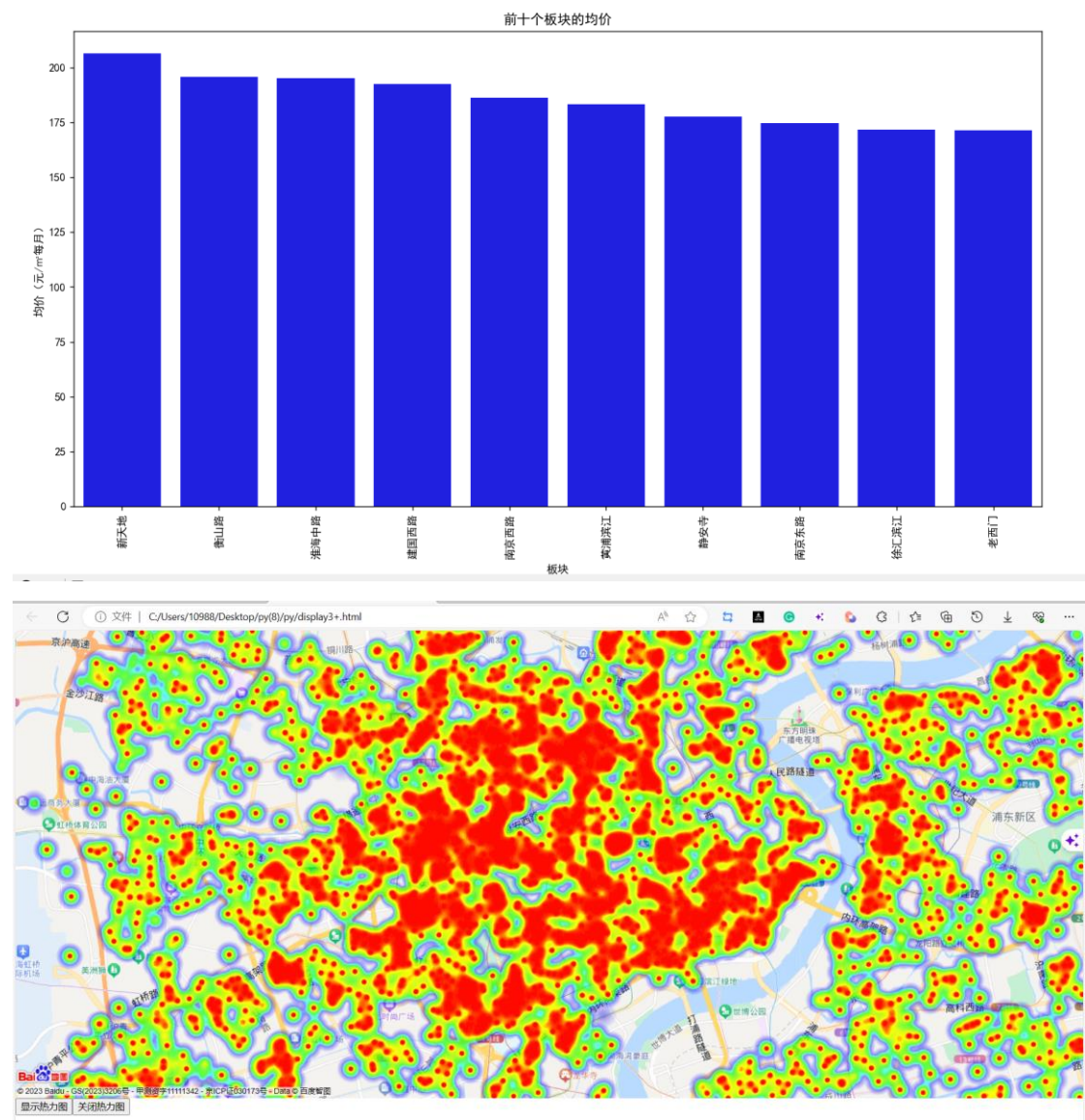
3.3 各城市不同板块均价

北京:

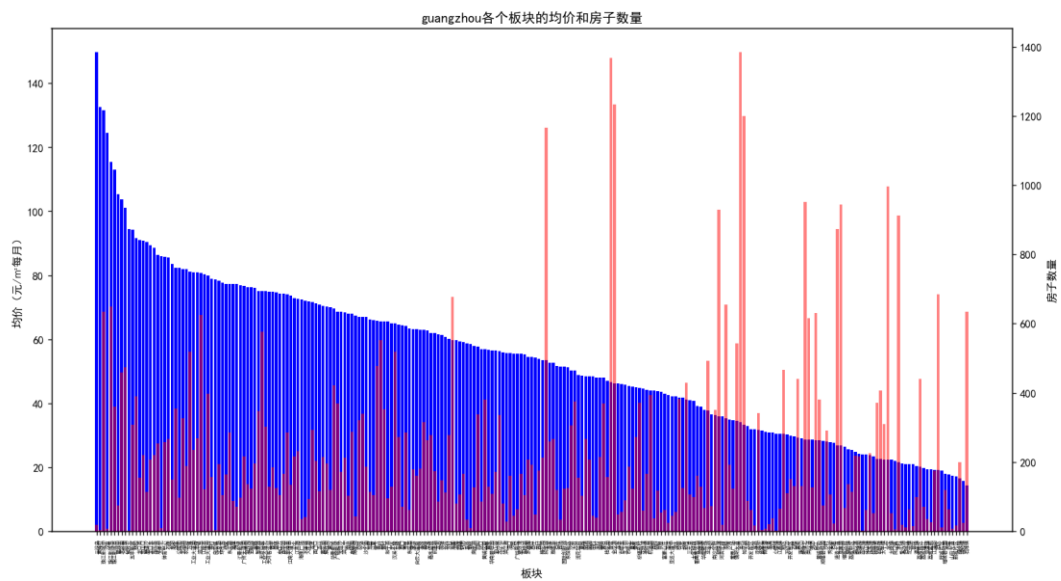


上海:

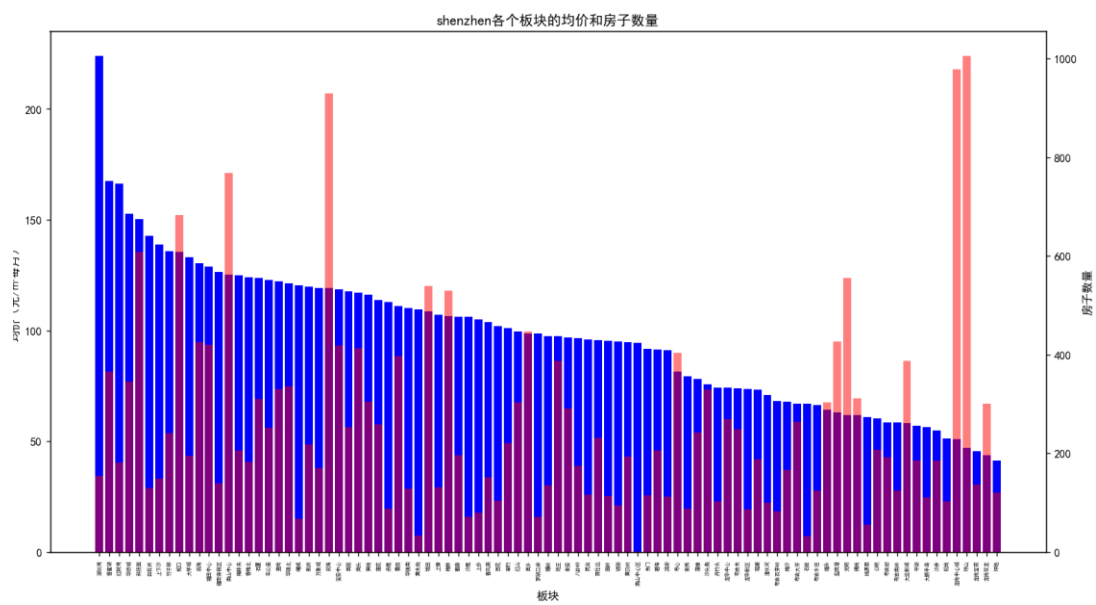




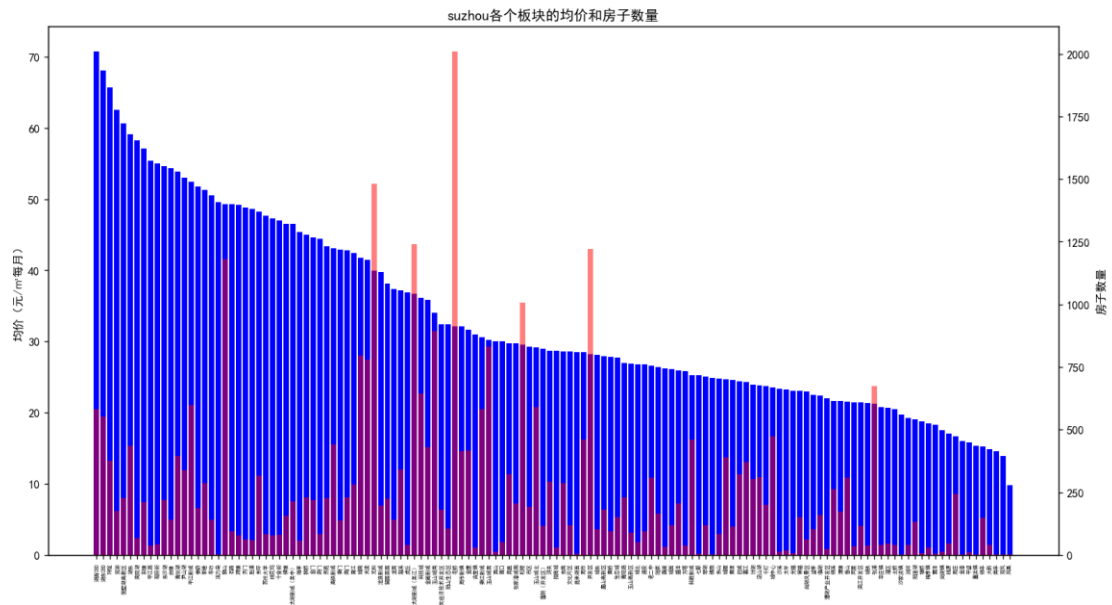
广州：



深圳：



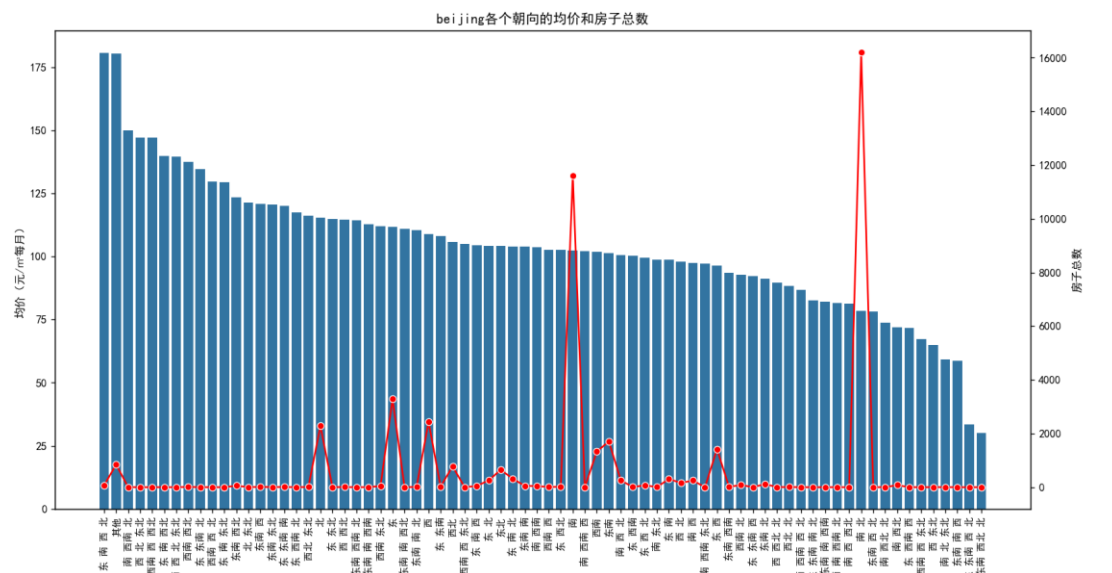
苏州：



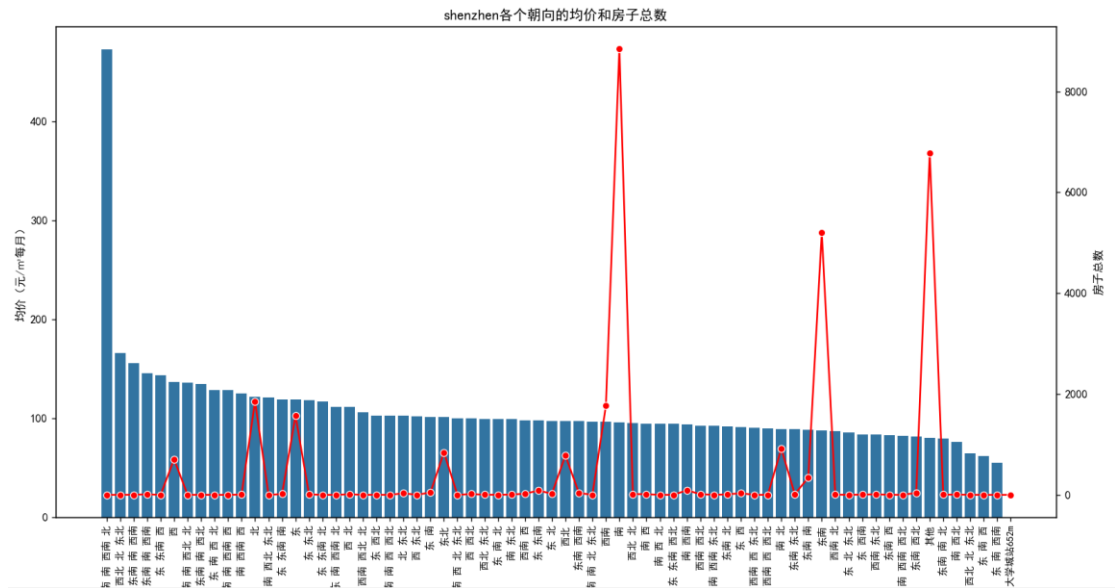
从上述图表可以看出各个板块的均价以及在租房子数量，以此可以更好的做租房决策等。从租房房价热力图中可以更具体地看出板块位置。

3.4 不同朝向对租金的影响

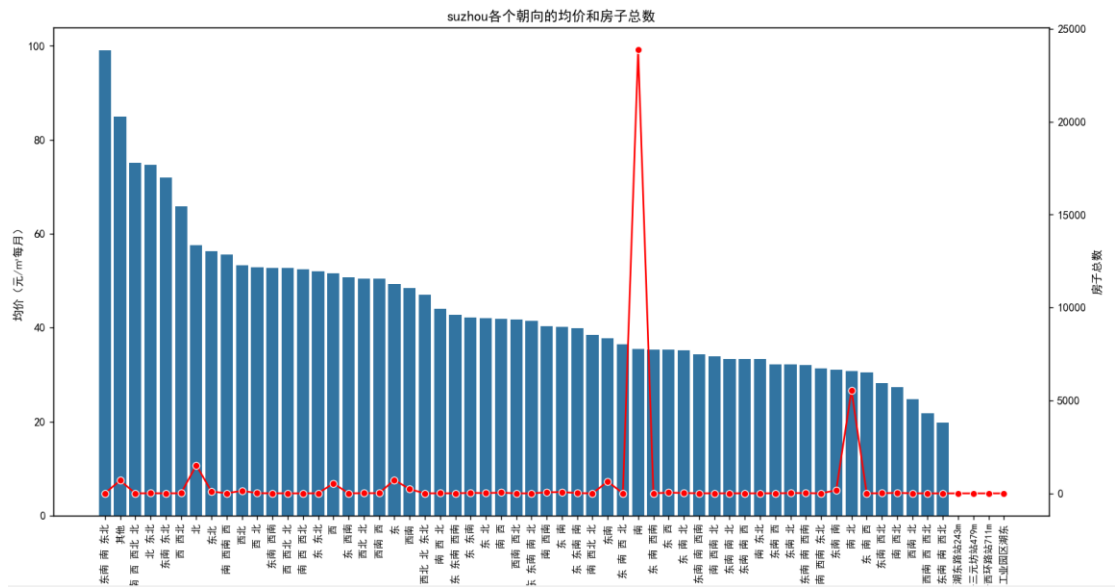
北京：



上海：



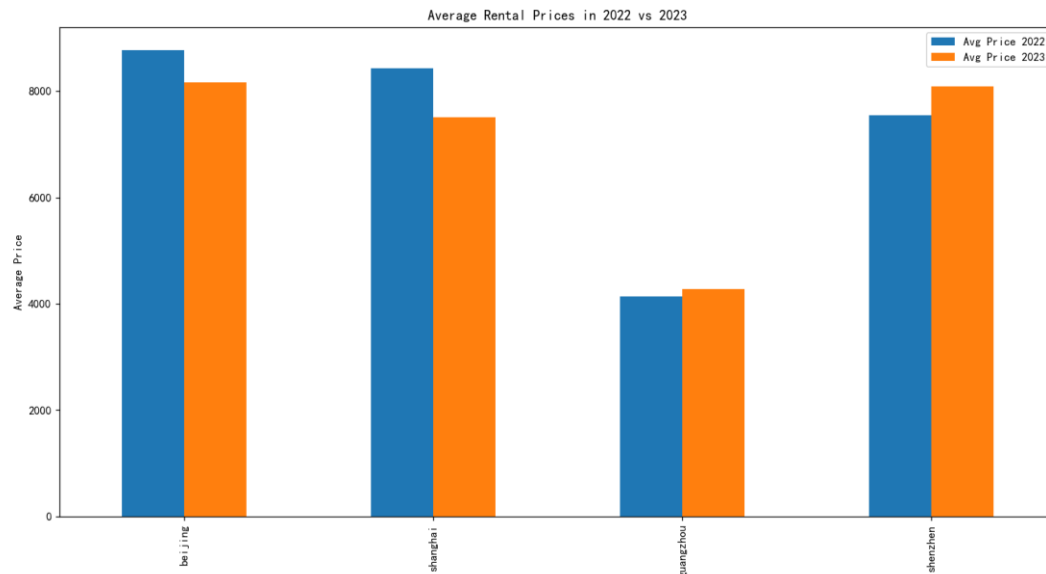
苏州：



首先,朝南的房子在五个城市都是最多的。也许是因为最多的原因,汇集了较多的低端租房,朝南房价最低。广州和深圳是朝西最贵,北京上海和苏州都是朝北最贵。

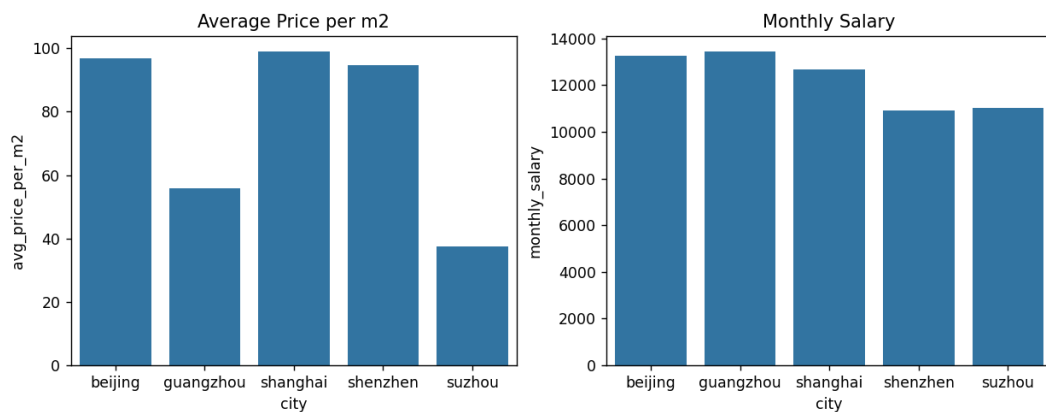
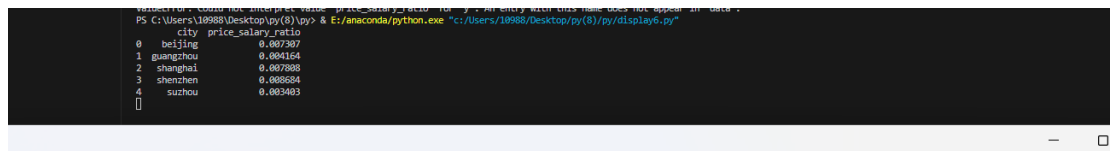
可能的原因:广州和深圳,冬季相对较暖,阳光充足,因此西晒可能被认为更为舒适。北京和上海,冬季相对较冷,而北向的房屋可能更容易保持温暖,因此更受青睐。

3.5 2023 与 2022 对比



平均租房价格 2022 到 2023 年，北京上海下降，上海下降较多，广州深圳上升。

3.6 工资与租房价格



从租房价格与薪资比例来看，深圳租房负担最重，北京上海其次，广州和苏州负担较小。

4 报告总结

首先我依次完成了作业要求，并给出了对可视化图表作出合理分析。值得注意的是，作业要求中，统计最高价和最低价似乎没有太大的意义。

其次我在要求基础上，获取了板块经纬坐标并绘制了热力图，更好地展示分析了板块均价数据。

目前数据分析较为粗糙，可以更细致地处理并分析。比如分析朝向可以控制变量进行分析，限制在一个板块内等。

但总的来说，我完整完成了任务，运用了爬虫，数据处理，可视化知识与函数库。