

第一次作业——武汉房屋数据分析报告

霍超

2024-10-31

你的主要发现

1. 发现 1 武汉房屋均价为 15148，房屋总价平均为 155.9w
2. 发现 2 武汉房屋的特点：精装为主，多为南向，板楼为主
3. 发现 3 武汉房屋白沙洲的房子供应最多，其次是盘龙城、四新、光谷东等，其中 CBD 西北湖、黄埔永清、中北路的房价最高

数据介绍

本报告链家数据获取方式如下：

报告人在 2023 年 9 月 12 日获取了链家武汉二手房网站数据。

- 链家二手房网站默认显示 100 页，每页 30 套房产，因此本数据包括 3000 套房产信息；
- 数据包括了页面可见部分的文本信息，具体字段及说明见作业说明。

说明：数据仅用于教学；由于不清楚链家数据的展示规则，因此数据可能并不是武汉二手房市场的随机抽样，结论很可能有很大的偏差，甚至可能是错误的。

...

数据概览

数据表（data）共包括等个变量,共行。表的前 10 行示例如下：

```
> head(data,10)
```

	property_name	property_region	price_ttl	price_sqm	bedrooms	livingrooms	building_area
1	南湖名都A区	南湖沃尔玛	237.0	18709	3	1	126.68
2	万科紫悦湾	光谷东	127.0	14613	3	2	86.91
3	东立国际	二七	75.0	15968	1	1	46.97
4	新都汇	光谷广场	188.0	15702	3	2	119.73
5	保利城一期	团结大道	182.0	17509	3	2	103.95
6	加州橘郡	庙山	122.0	10376	3	2	117.59
7	省建筑五公司西区	光谷广场	99.0	12346	2	1	80.19
8	保利上城东区	白沙洲	193.8	16336	3	2	118.64
9	石化大院	中南丁字桥	325.0	32631	4	1	99.60
10	阳光花园	杨汊湖	192.0	17403	3	2	110.33

	directions1	directions2	decoration	property_t_height	property_height	property_style	followers
1	南	北	精装	17	中	塔楼	3
2	南		精装	28	中	板楼	1
3	南		简装	18	中	塔楼	3
4	北	东	精装	32	中高	塔楼	2
5	东南		简装	34	中	板塔结合	3
6	南	北	精装	34	中低	板楼	1
7	南		简装	7	低	板楼	0
8	南	北	其他	34	中	板塔结合	0
9	南	北	简装	5	中低	板楼	2
10	南	北	其他	7	低	板楼	0

	near_subway	if_2y	has_key	vr
1	近地铁	<NA>	随时看房	<NA>
2	<NA>	房本满两年	随时看房	VR看装修
3	近地铁	<NA>	随时看房	<NA>
4	近地铁	房本满两年	随时看房	<NA>
5	<NA>	房本满两年	随时看房	VR看装修
6	<NA>	房本满两年	随时看房	<NA>
7	近地铁	<NA>	随时看房	VR看装修
8	近地铁	房本满两年	随时看房	<NA>
9	近地铁	<NA>	随时看房	<NA>
10	近地铁	房本满两年	随时看房	<NA>

各变量的简短信息:

```
***{r}
glimpse(data)
columns: 18
$ property_name      <chr> "南湖名都A区", "万科紫悦湾", "东立国际", "新都汇", "保利城一期", "...
$ property_region    <chr> "南湖沃尔玛", "光谷东", "二七", "光谷广场", "团结大道", "庙山", "...
$ price_ttl          <dbl> 237.0, 127.0, 75.0, 188.0, 182.0, 122.0, 99.0, 193.8, 325.0, 192.0, ...
$ price_sqm          <int> 18709, 14613, 15968, 15702, 17509, 10376, 12346, 16336, 32631, 174...
$ bedrooms           <int> 3, 3, 1, 3, 3, 3, 2, 3, 4, 3, 5, 3, 4, 3, 3, 2, 3, 4, 3, 3, 3, 3, ...
$ livingrooms        <int> 1, 2, 1, 2, 2, 2, 1, 2, 1, 2, 2, 2, 2, 1, 2, 2, 2, 2, 2, 2, 1, 2, ...
$ building_area       <dbl> 126.68, 86.91, 46.97, 119.73, 103.95, 117.59, 80.19, 118.64, 99.60, ...
$ directions1        <chr> "南", "南", "南", "北", "东南", "南", "南", "南", "南", "南", "南", ...
$ directions2        <chr> "北", "", "", "东", "", "北", "", "北", "北", "北", "北", "西...
$ decoration         <chr> "精装", "精装", "简装", "精装", "简装", "精装", "简装", "其他", "...
$ property_t_height   <int> 17, 28, 18, 32, 34, 34, 7, 34, 5, 7, 25, 32, 8, 31, 44, 19, 7, 8, ...
$ property_height     <chr> "中", "中", "低", "高", "中", "低", "低", "中", "低", "低", "高", ...
$ property_style      <chr> "塔楼", "板楼", "塔楼", "塔楼", "板塔结合", "板楼", "板楼", "板塔...
$ followers           <int> 3, 1, 3, 2, 3, 1, 0, 0, 2, 0, 0, 0, 10, 0, 0, 1, 0, 1, 1, 0, 2, 3, ...
$ near_subway        <chr> "近地铁", NA, "近地铁", "近地铁", NA, NA, "近地铁", "近地铁", "近...
$ if_2y              <chr> NA, "房本满两年", NA, "房本满两年", "房本满两年", "房本满两年", NA...
$ has_key             <chr> "随时看房", "随时看房", "随时看房", "随时看房", "随时看房", "随时...
$ vr                 <chr> NA, "VR看装修", NA, NA, "VR看装修", NA, "VR看装修", NA, NA, NA, NA...
```

各变量的简短统计:

property_name	property_region	price_ttl	price_sqm	bedrooms
Length:3000	Length:3000	Min. : 10.6	Min. : 1771	Min. :1.000
Class :character	Class :character	1st Qu.: 95.0	1st Qu.:10799	1st Qu.:2.000
Mode :character	Mode :character	Median : 137.0	Median :14404	Median :3.000
		Mean : 155.9	Mean :15148	Mean :2.695
		3rd Qu.: 188.0	3rd Qu.:18211	3rd Qu.:3.000
		Max. :1380.0	Max. :44656	Max. :7.000
livingrooms	building_area	directions1	directions2	decoration
Min. :0.000	Min. : 22.77	Length:3000	Length:3000	Length:3000
1st Qu.:1.000	1st Qu.: 84.92	Class :character	Class :character	Class :character
Median :2.000	Median : 95.55	Mode :character	Mode :character	Mode :character
Mean :1.709	Mean :100.87			
3rd Qu.:2.000	3rd Qu.:117.68			
Max. :4.000	Max. :588.66			
property_t_height	property_height	property_style	followers	near_subway
Min. : 2.00	Length:3000	Length:3000	Min. : 0.000	Length:3000
1st Qu.:11.00	Class :character	Class :character	1st Qu.: 1.000	Class :character
Median :27.00	Mode :character	Mode :character	Median : 3.000	Mode :character
Mean :24.22			Mean : 6.614	
3rd Qu.:33.00			3rd Qu.: 6.000	
Max. :62.00			Max. :262.000	
if_2y	has_key	vr		
Length:3000	Length:3000	Length:3000		
Class :character	Class :character	Class :character		
Mode :character	Mode :character	Mode :character		

可以看到:

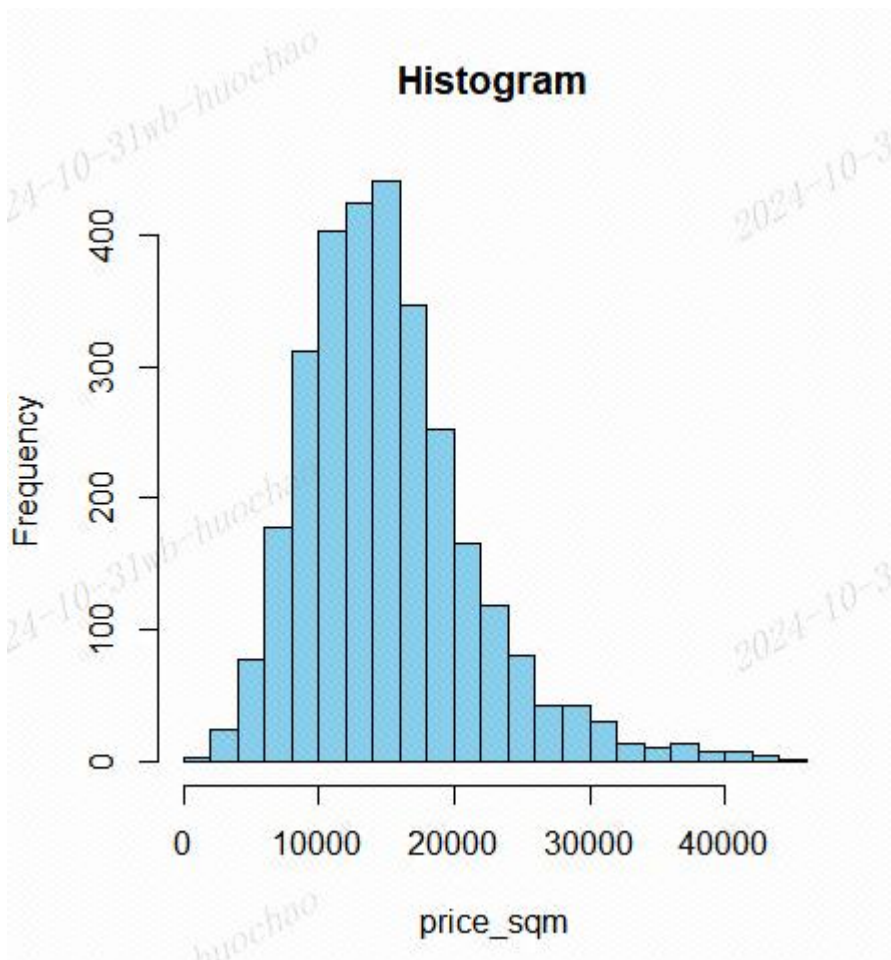
- 直观结论 1 房屋平均价格为 155.9 万元，平均单价 15148 元，建筑面积平均值为 100.87 平方米
- 直观结论 2 房屋总价范围为 10.6-1380 万元，每平方米单价范围 1771-44656 元
- 直观结论 3 卧室主要是 2-3 居室，购买的楼层平均为 6 层

探索性分析

价格描述与图形

发现：最低房价为 1771，最高房价为 44656，平均价格为 15148，房屋中位数价格为 14404，标准差为 6323.175

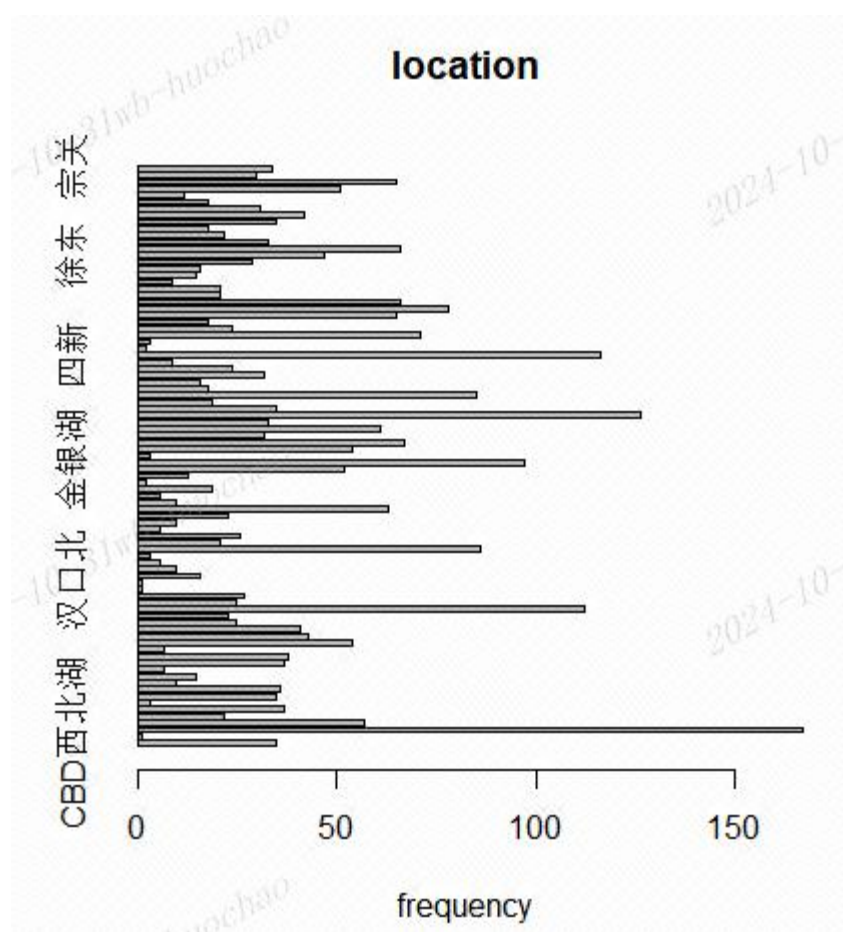
- 发现 1: 房屋价格中间高，两边低，呈现正太分布左偏主要价格区间在 0-25000
- 发现 2: 房屋价格主要集中在在 1771-25000 之间，超过 32000 以上的房屋比较少

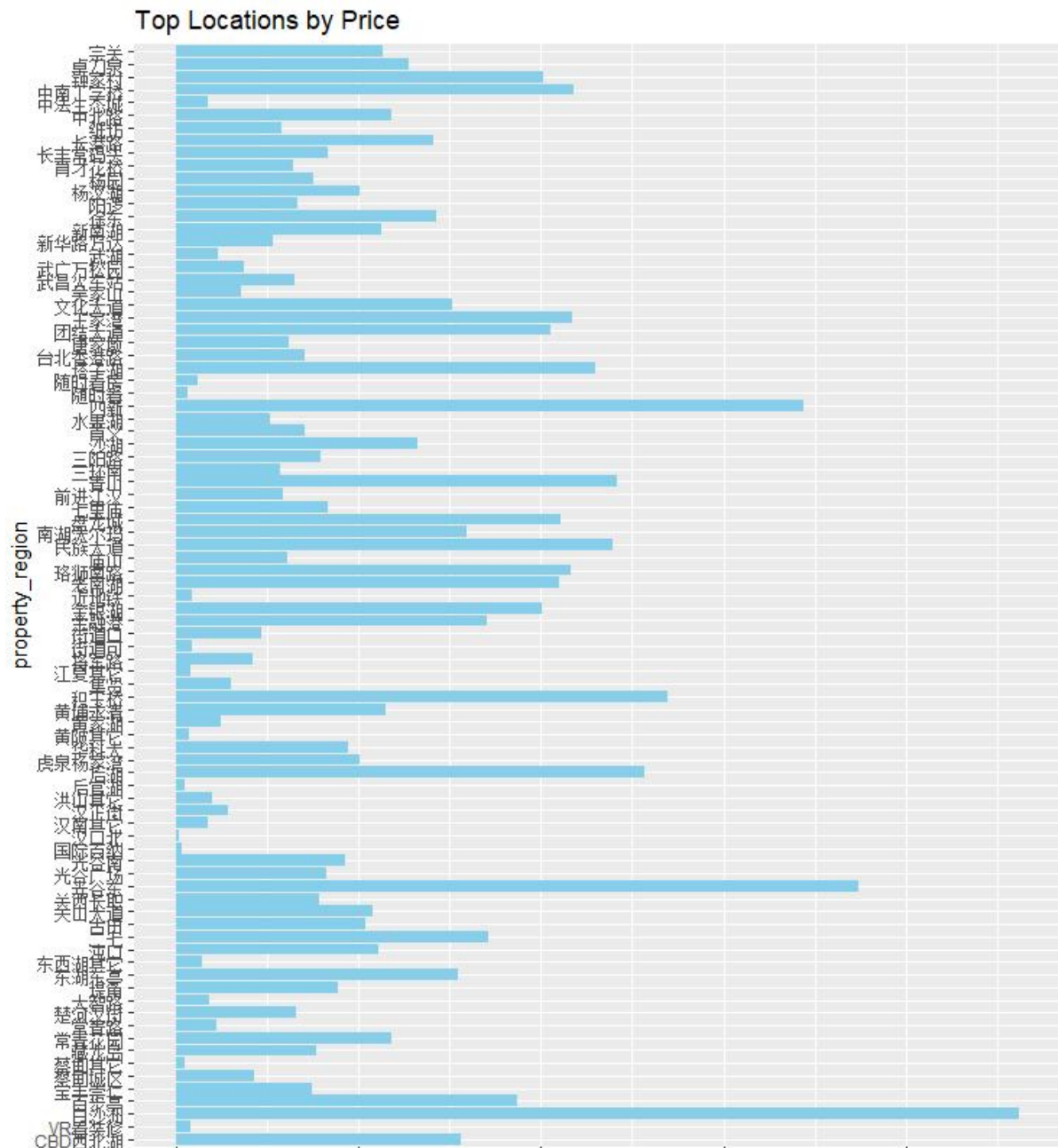


区位的数值描述与图形

发现：房屋主要的位置在 CBD 西北湖、汉口北、金银湖、四新、徐东、宗关

- 发现 1 白沙洲的房子供应最多，其次是盘龙城、四新、光谷东等
- 发现 2 供应房子多的地方大部分位于郊区、也就是市区内土地供应优先，房屋都往郊区建设

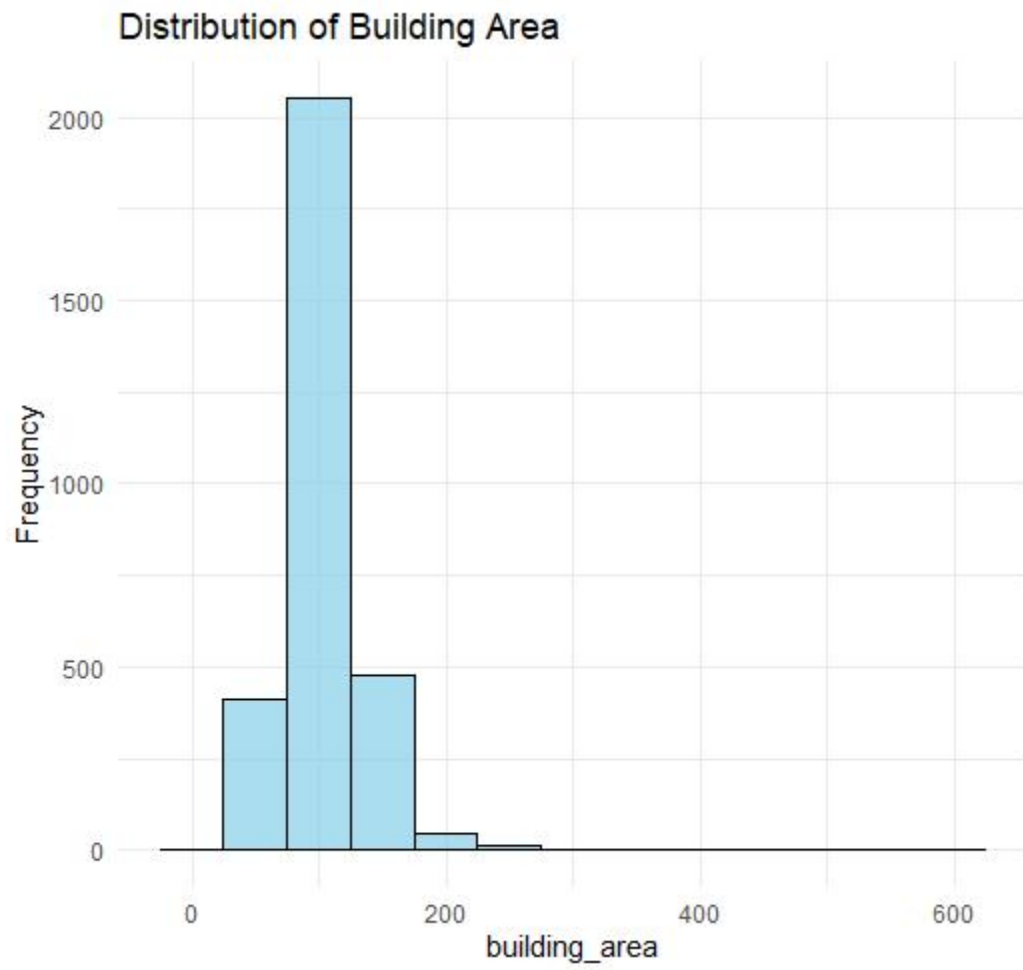


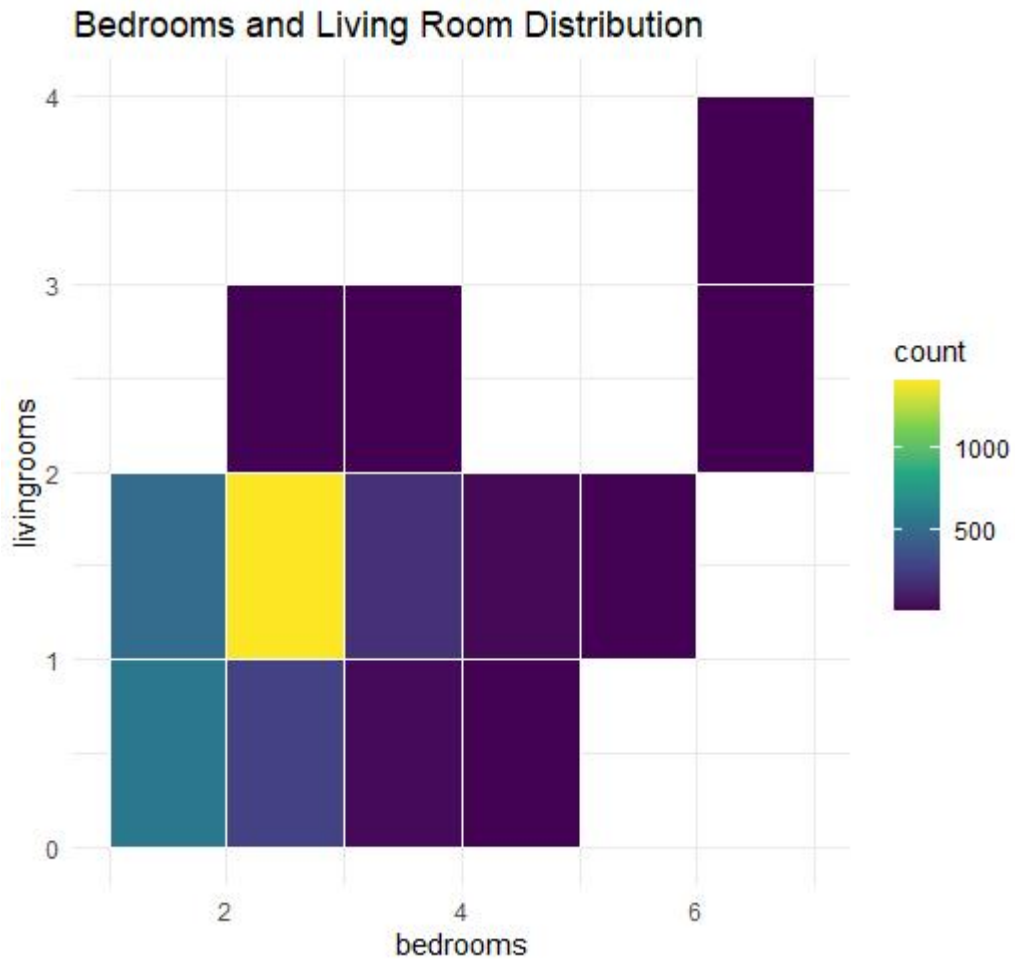


建筑面积、客厅数和房间数的数值描述与图形

发现：建筑面积主要集中在 0-200 平方之间，超过 200 平方的房屋很少，武汉市房屋的建筑面积为 100 平方左右最为常见

- 发现 1 武汉房屋 3 居室 1 厅或 2 厅居多
- 发现 2 房屋普遍在 4 居室以内，客厅数普遍在 2 厅以内，即房屋的房间数和客厅数大多数不会很多

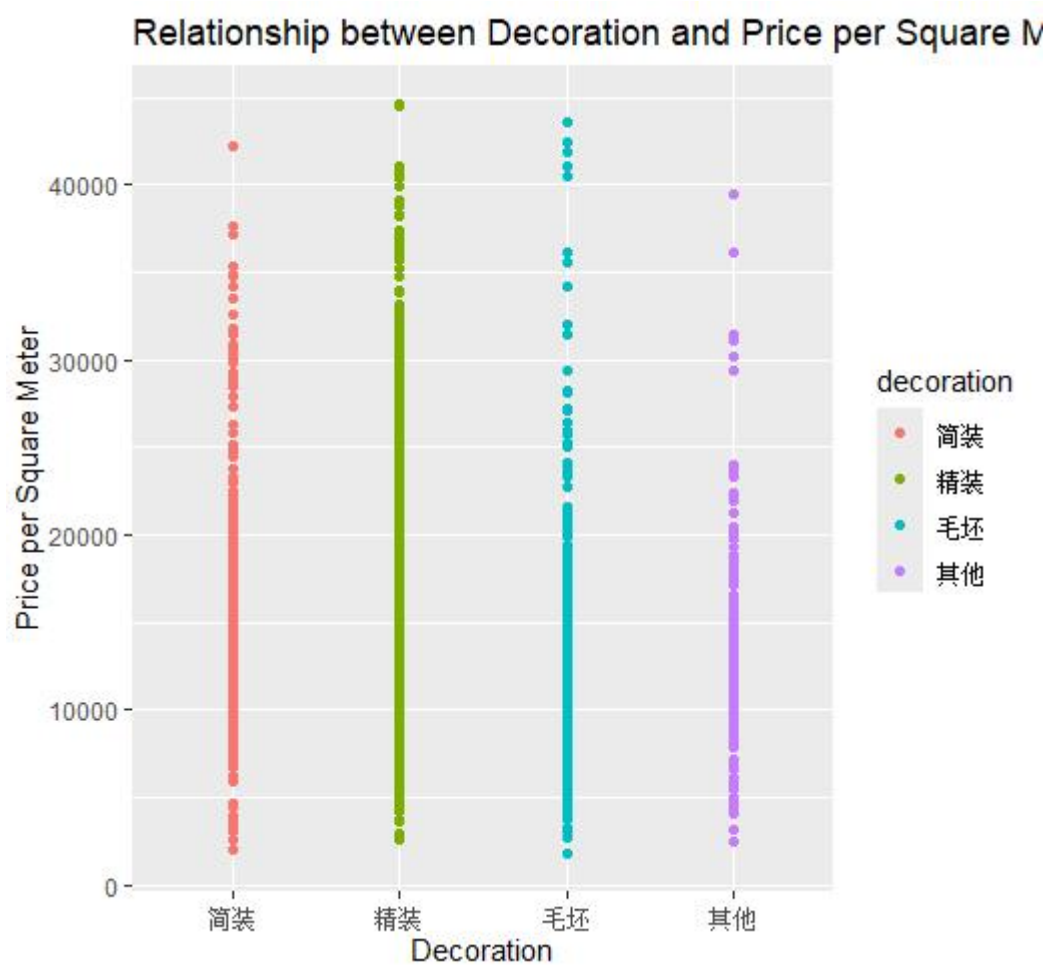




探索问题 1 装修和价格的关系，分析装修状况是否影响单价

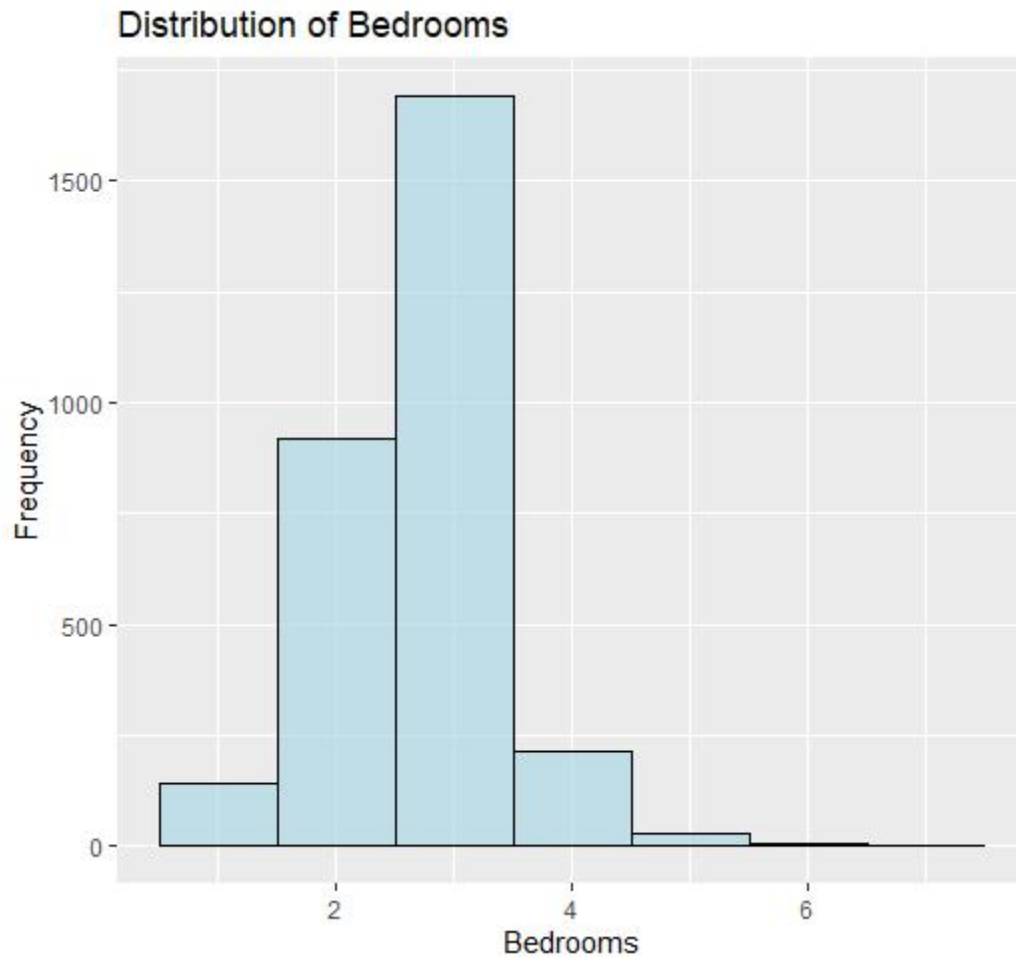
发现：装修对价格有一定影响，但影响不是很大

- 发现 1 武汉房屋精装较多其次是毛坯，简装较少
- 发现 2 精装的房屋价格确实比简装和毛坯稍贵一点



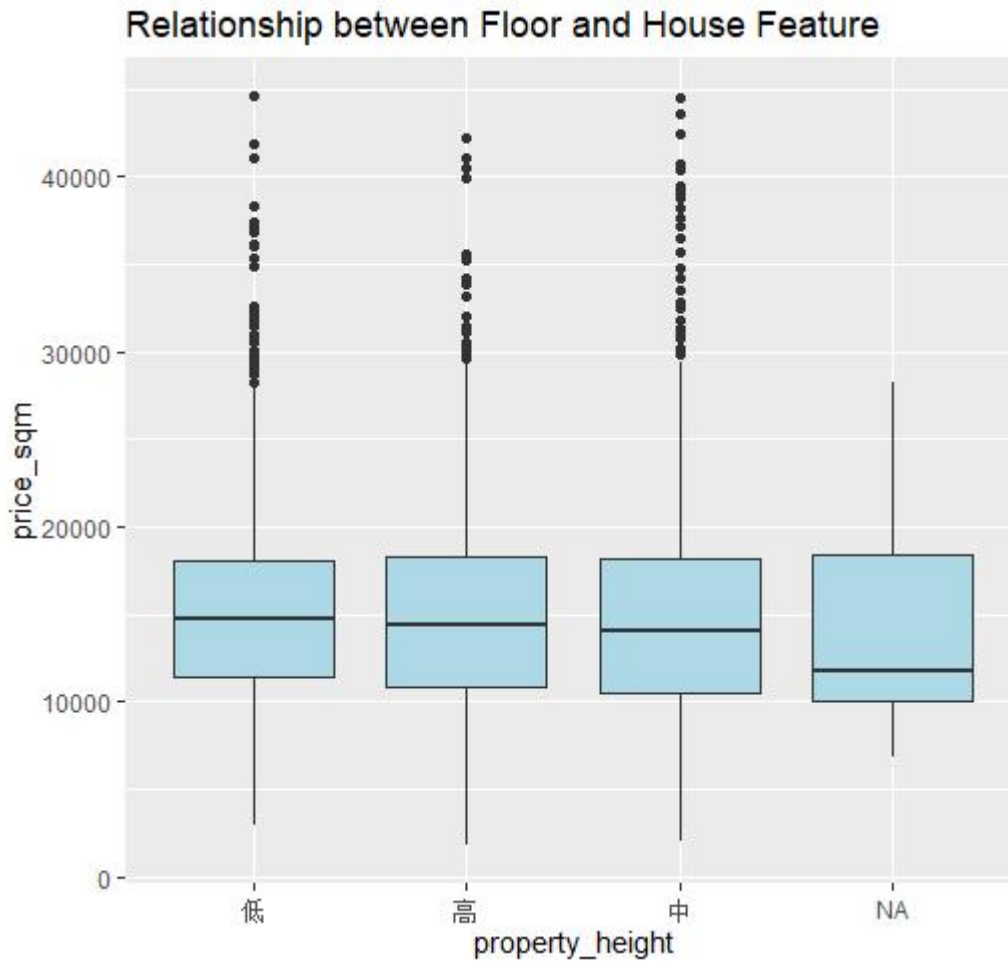
探索问题 2 武汉房屋卧室数量的分布

发现：3 居室最常见，其次是 2 居室，很少有超过 6 居室的房屋



探索问题 3 楼层位置和房屋单价的关系

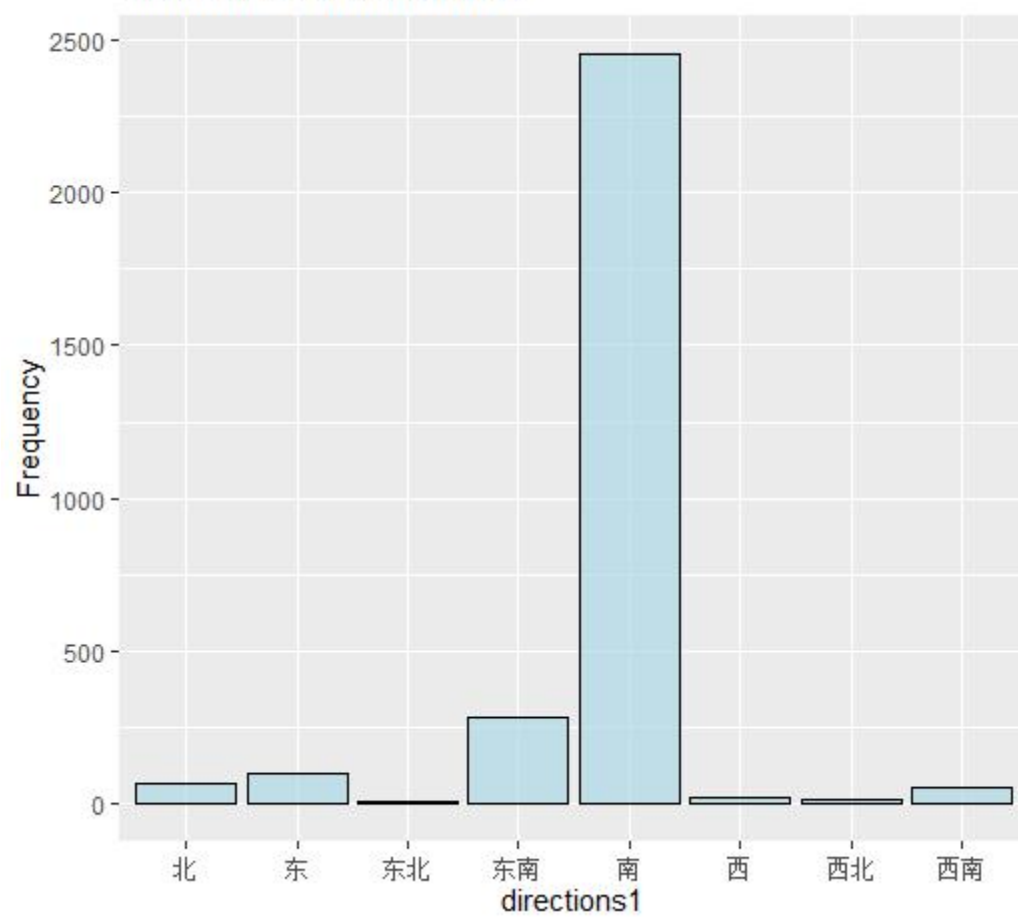
发现：低楼层的房屋价格确实要低一点，中高层的房屋价格稍高一点，但并没有高很多

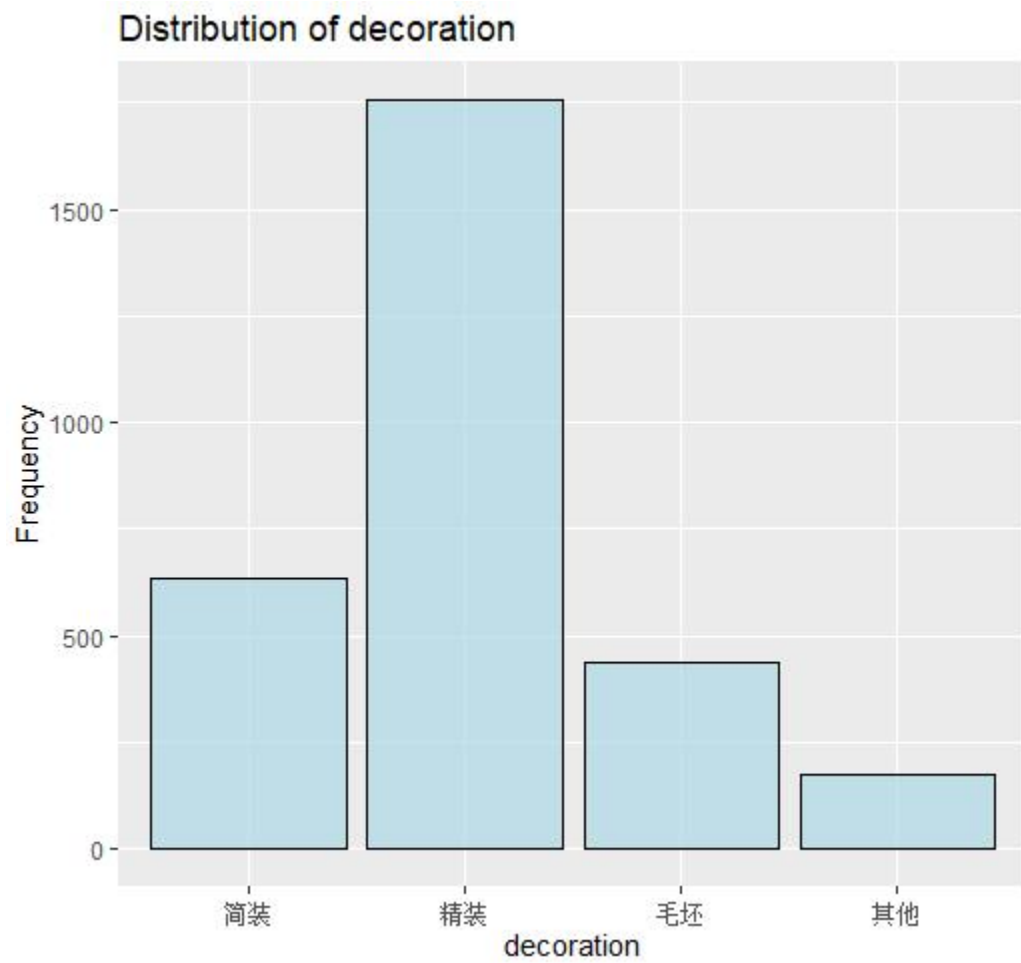


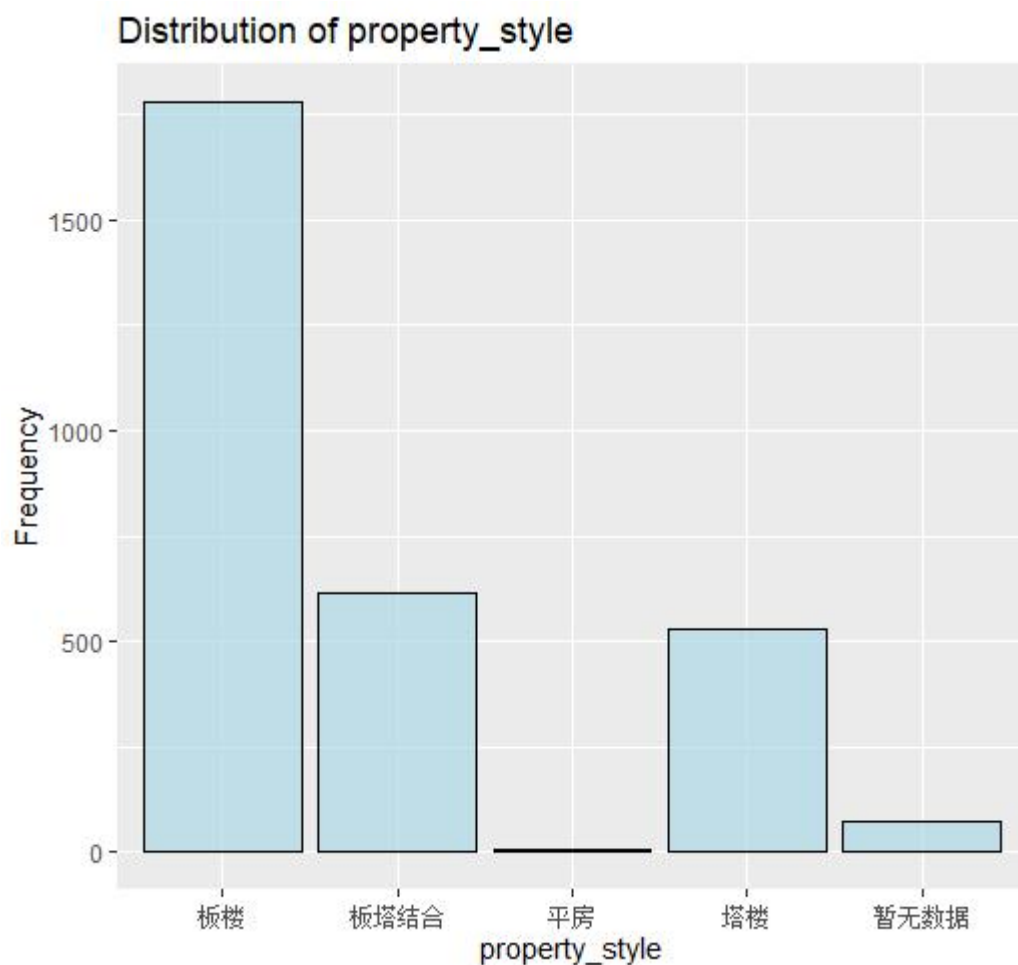
探索问题 4 武汉房屋楼型、装修风格、朝向的分布

发现 1 武汉房屋板楼居多，其次是塔楼和板塔结合，基本无平房
发现 2 武汉房屋精装居多，简装和毛坯较少
发现 3 武汉房屋集中于朝南，其次是东南向，其他朝向很少

Distribution of directions1







发现总结

用 1-3 段话总结你的发现。

- 1、武汉房价比较高的区域是 CBD 西北湖、黄埔永清、中北路、中南丁字桥，房价比较低的区域是阳逻、汉南其他，也就是中心城区的价格贵，偏远郊区价格便宜
- 2、武汉房屋主要还是朝南向，精装为主、有部分简装和毛坯，并且是板楼居多，可能与武汉天气有关系
- 3、武汉房屋价格总价范围为 10.6-1380 万元，平均价格 155.9 万元；单价范围 1771-44656 元，平均 15148 元，卧室数范围为 1-7,2-3 居室最为常见，靠近地铁的房源更受欢迎