

2023MEM第一次作业

陈杰 2023281051009

2023-10-20

你的主要发现

- 1. 发现1
- 2. 发现2
- 3. 发现3

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 (<https://wh.lianjia.com/ershoufang/>)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 (lj)共包括property_name, property_region, price_ttl, price_sqm, bedrooms, livingrooms, building_area, directions1, directions2, decoration, property_t_height, property_height, property_style, followers, near_subway, if_2y, has_key, vr等18个变量,共3000行。表的前10行示例如下：

各变量的简短信息：

## Rows: 3,000	
## Columns: 18	
## \$ property_name	<chr> "南湖名都A区", "万科紫悦湾", "东立国际", "新都汇", "...
## \$ property_region	<chr> "南湖沃尔玛", "光谷东", "二七", "光谷广场", "团结大...
## \$ price_ttl	<dbl> 237.0, 127.0, 75.0, 188.0, 182.0, 122.0, 99.0, 193.8...
## \$ price_sqm	<dbl> 18709, 14613, 15968, 15702, 17509, 10376, 12346, 163...
## \$ bedrooms	<dbl> 3, 3, 1, 3, 3, 3, 2, 3, 4, 3, 5, 3, 4, 3, 3, 2, 3, 4...
## \$ livingrooms	<dbl> 1, 2, 1, 2, 2, 2, 1, 2, 1, 2, 2, 2, 2, 1, 2, 2, 2, 2...
## \$ building_area	<dbl> 126.68, 86.91, 46.97, 119.73, 103.95, 117.59, 80.19, ...
## \$ directions1	<chr> "南", "南", "南", "北", "东南", "南", "南", "南", "南", "...
## \$ directions2	<chr> "北", NA, NA, "东", NA, "北", NA, "北", "北", "北", ...
## \$ decoration	<chr> "精装", "精装", "简装", "精装", "简装", "精装", "简...
## \$ property_t_height	<dbl> 17, 28, 18, 32, 34, 34, 7, 34, 5, 7, 25, 32, 8, 31, ...
## \$ property_height	<chr> "中", "中", "低", "高", "中", "低", "低", "中", "低"...
## \$ property_style	<chr> "塔楼", "板楼", "塔楼", "塔楼", "板塔结合", "板楼", ...
## \$ followers	<dbl> 3, 1, 3, 2, 3, 1, 0, 0, 2, 0, 0, 0, 10, 0, 0, 1, 0, ...
## \$ near_subway	<chr> "近地铁", NA, "近地铁", "近地铁", NA, NA, "近地铁", ...
## \$ if_2y	<chr> NA, "房本满两年", NA, "房本满两年", "房本满两年", "...
## \$ has_key	<chr> "随时看房", "随时看房", "随时看房", "随时看房", "随...
## \$ vr	<chr> NA, "VR看装修", NA, NA, "VR看装修", NA, "VR看装修", ...

各变量的简短统计：

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

可以看到：

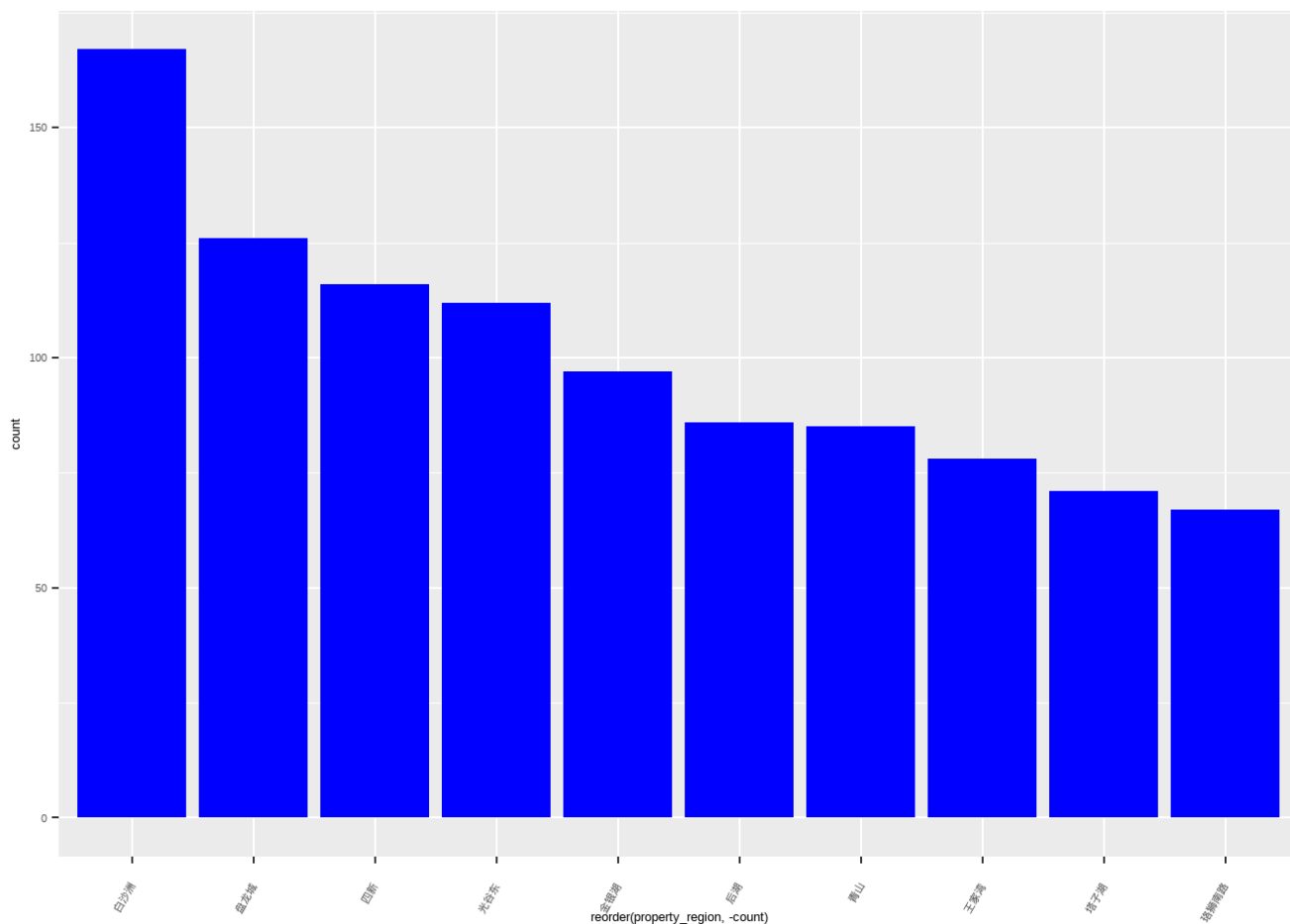
- 直观结论1 武汉二手房总价均值为155.9万元，中位数为137万元；总价金额主要集中在95万元至188万元区间内。
- 直观结论2 武汉二手房单价均值为15148元，中位数为14404元；单价主要集中在10799元至18211元区间内。
- 直观结论3 武汉二手房以2居室和3居室居多，占比超过50%；建筑面积集中在84.92m²至117.68m²区间内，平均建筑面积为100.87m²，中位数为95.55m²。

探索性分析

by_subway的数值描述与图形

发现：

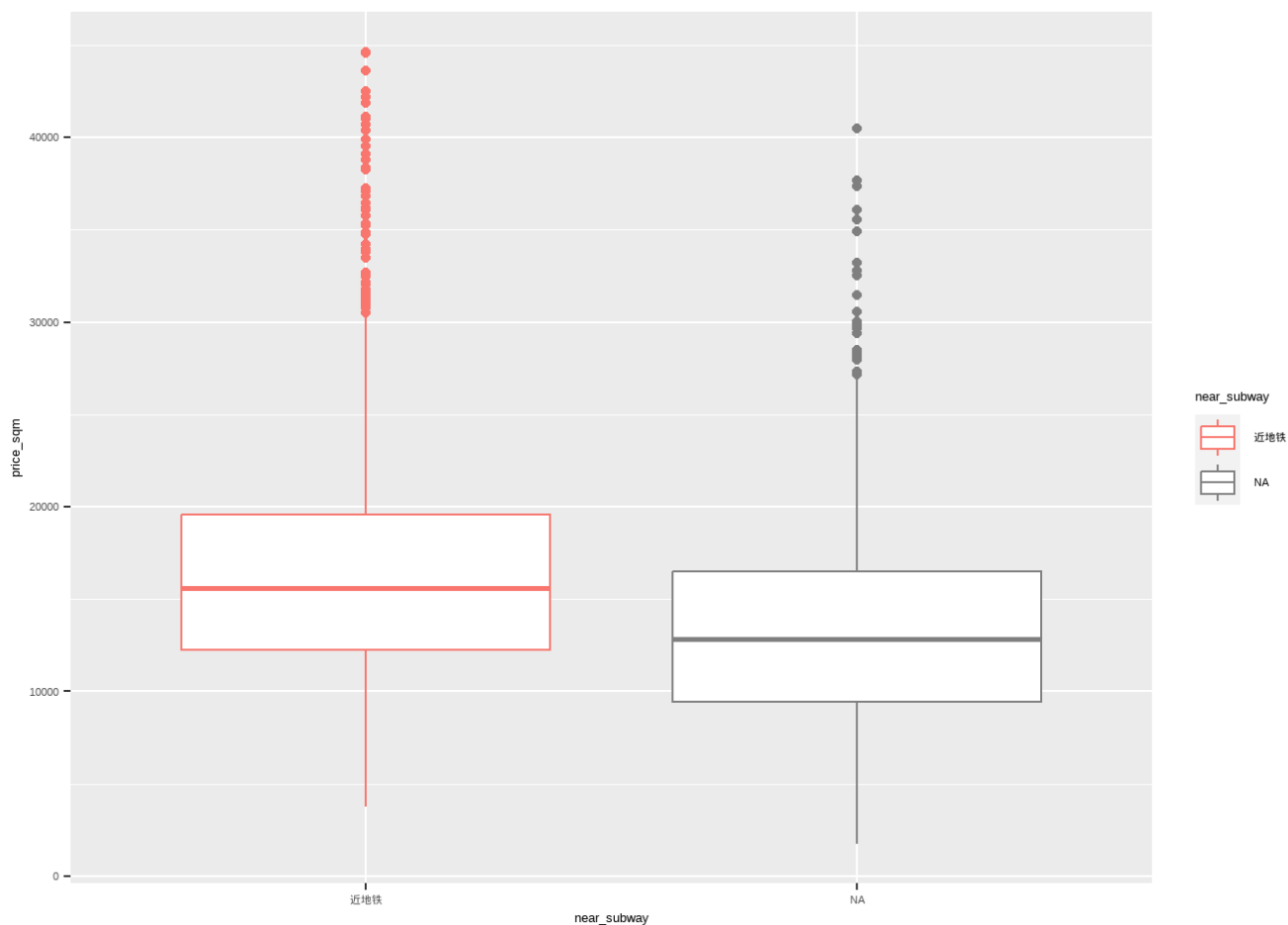
- 发现1:武汉二手房中，地铁沿线与非地铁沿线的房子比例大致为1：1。



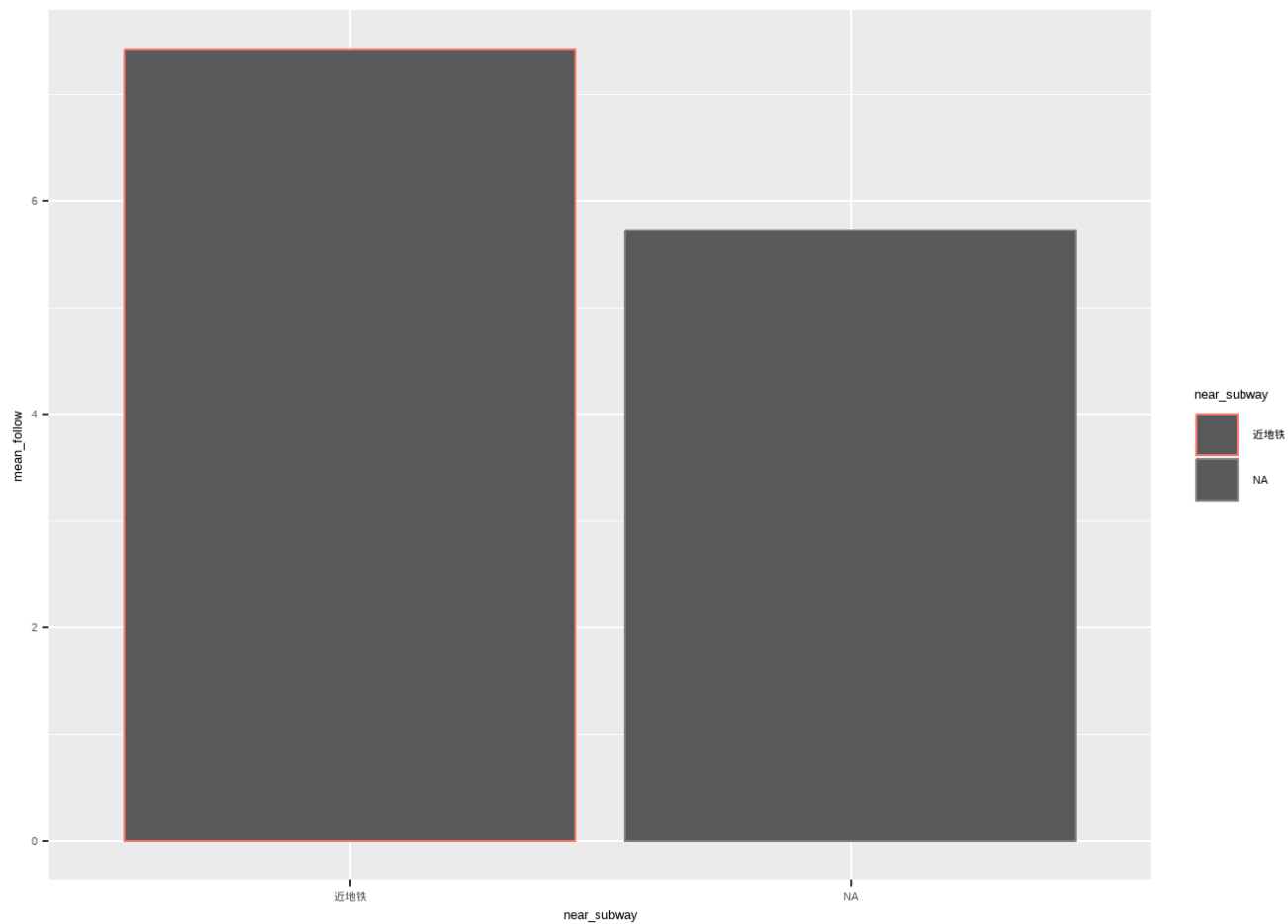
探索问题1：地铁是否影响二手房单价和关注度

发现：

- 发现1:地铁沿线房价明显高于非地铁沿线



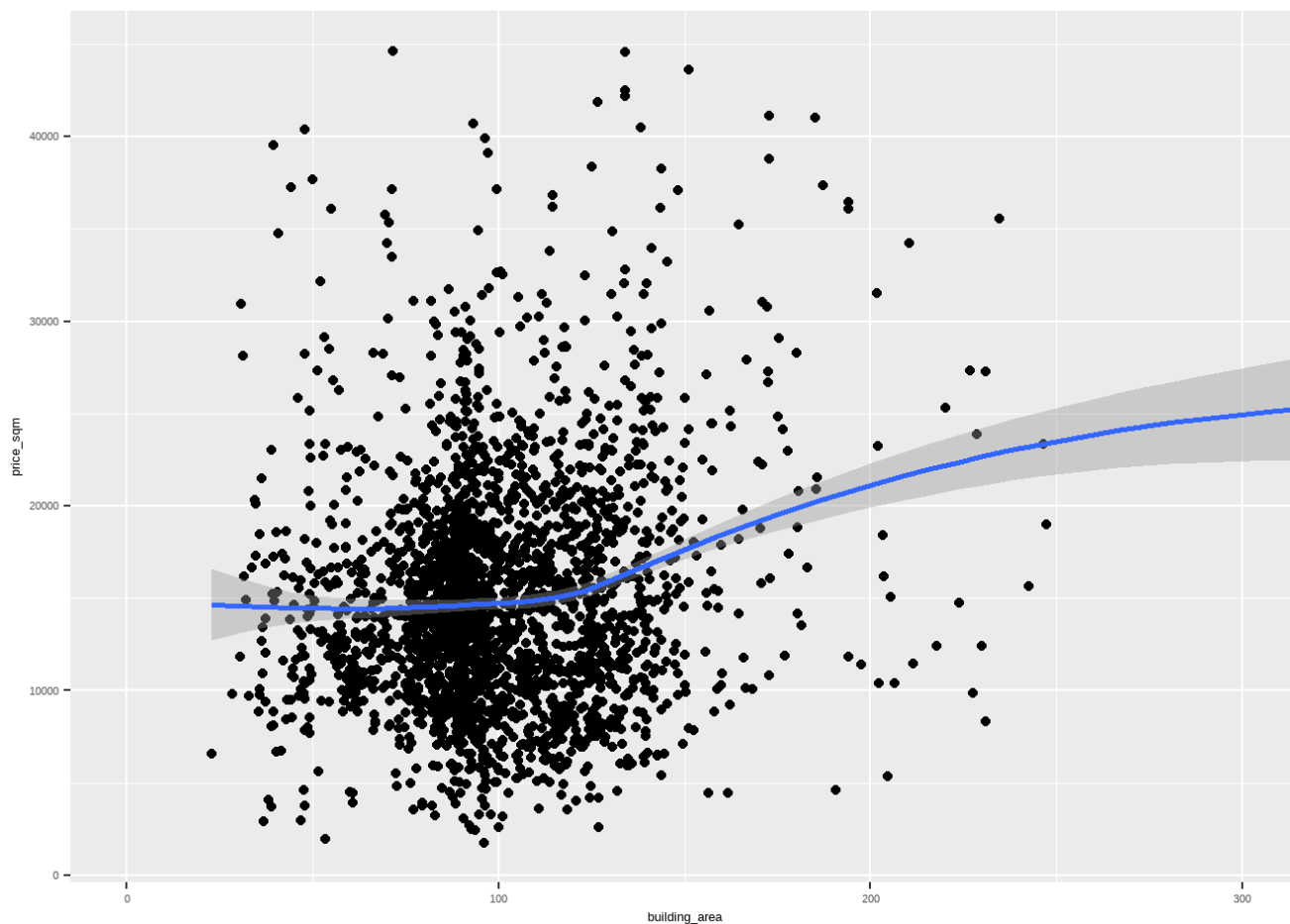
- 发现2:地铁沿线二手房关注度更高



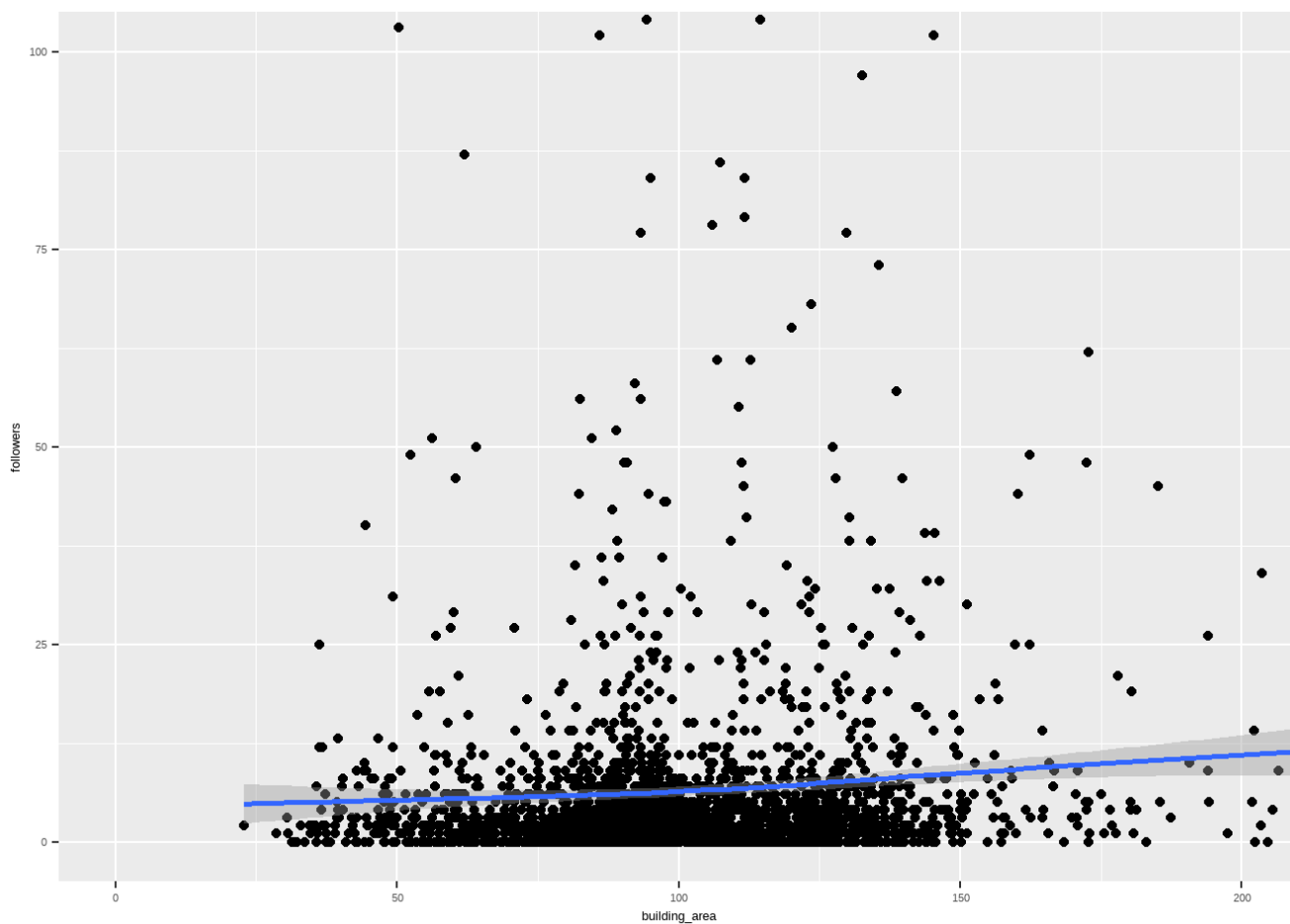
探索单价、面积、关注度直接的关系

发现:

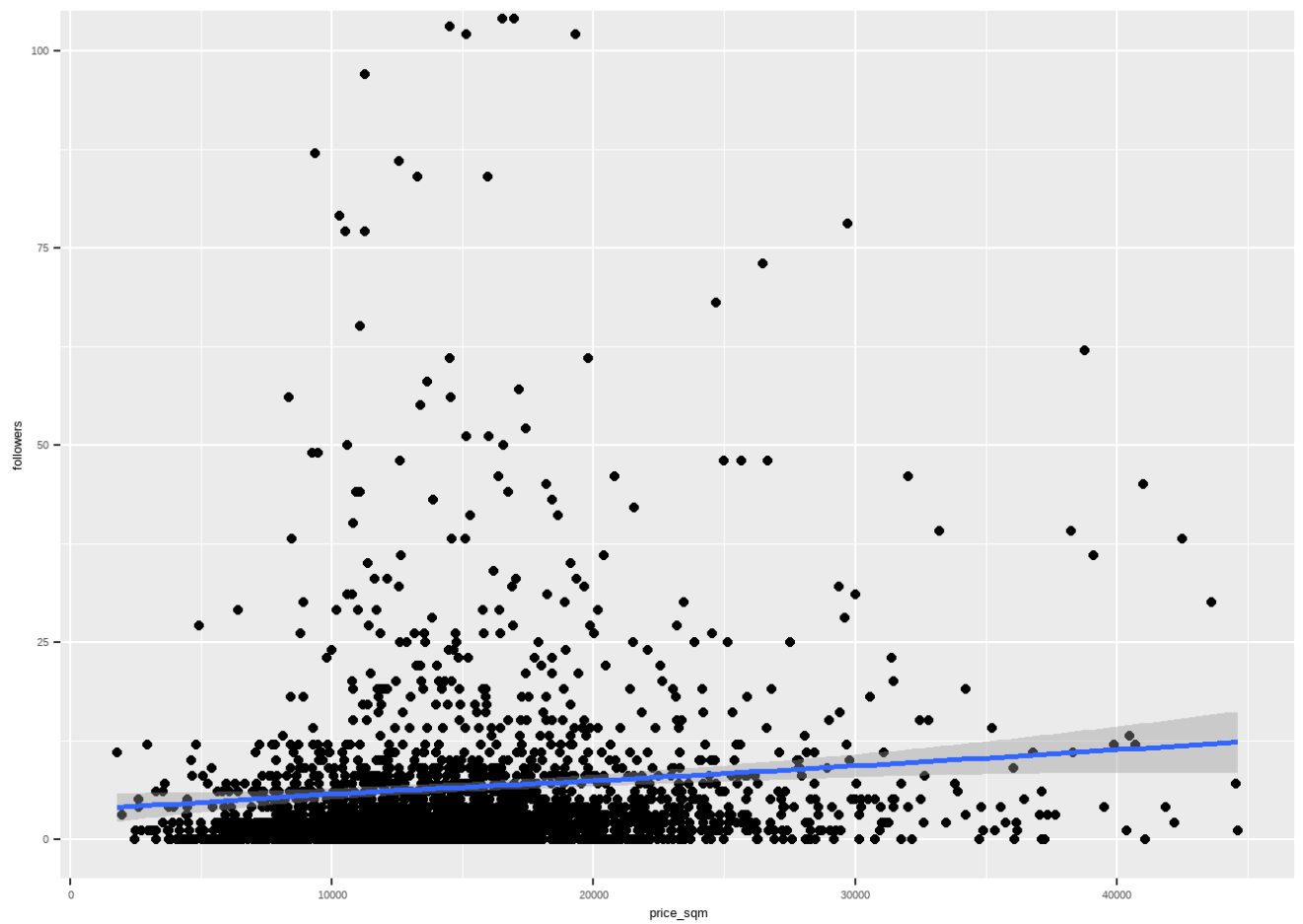
- 发现1:二手房单价整体随面积增大而升高



- 发现2:二手房关注度整体随面积增大而升高



- 发现3:二手房关注度整体随单价升高而升高



发现总结

1.地铁房具有明显市场价值，地铁房单价明显高于非地铁房。2.二手房市场供应按区域划分有明显的聚集点，且各聚集点在武汉三镇分布较分散。3.在正常数值范围内，二手房单价、面积、关注度三者呈正相关。