

1st_assignment_eda_Hoki_Chow

周青

2024-10-27

true

仓库中data目录下是来自某二手房网站有关武汉的3000套二手房的信息,请利用你掌握的描述性统计方法，对数据进行探索性数据分析，可以讨论的问题包括（但不限于）：价格特点、区位特点、建筑形式、以及这些变量的组合分析等。

你的主要发现

1. 发现1：建筑面积与房产总价对数呈正相关，区域发展潜力也影响房产总价，这些正相关关系反映房地产市场定价规律，为价格评估和预测提供依据。
2. 发现2：装修情况对房产单价有显著影响，精装修能提升单价，开发商和卖家可通过高质量精装修增加附加值，而楼层高度对单价影响不大。
3. 发现3：随时看房和 VR 看房的显著性体现了看房者对看房便利性的重视，前者可灵活安排时间、减少协调麻烦，后者利用现代科技让看房者在初步筛选阶段直观了解房屋、节省精力。靠近地铁在购房或租房决策中有显著影响，它代表的交通便利性可减少通勤成本和时间，表明大部分看房者可能是上班族或对交通便捷有要求的人群，房产周边交通设施吸引力大。

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(ij1)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：

武汉链家二手房

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
南湖名都A区 南湖	237.																
湖沃尔																	

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           <int> 18709, 14613, 15968, 15702, 17509, 10376, 12346, 163…
## $ bedrooms            <int> 3, 3, 1, 3, 3, 3, 2, 3, 4, 3, 5, 3, 4, 3, 3, 2, 3, 4…
## $ livingrooms         <int> 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", "…
## $ directions2        <chr> "北", "", "", "东", "", "北", "", "北", "北", "北", …
## $ decoration          <chr> "精装", "精装", "简装", "精装", "简装", "精装", "简…
## $ property_t_height   <int> 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           <int> 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
##
##
##
```

```
## 'data.frame': 3000 obs. of 18 variables:
## $ 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" ...
## $ price_ttl : num 237 127 75 188 182 ...
## $ price_sqm : int 18709 14613 15968 15702 17509 10376 12346 16336 32631 17403 ...
## $ bedrooms : int 3 3 1 3 3 3 2 3 4 3 ...
## $ livingrooms : int 1 2 1 2 2 2 1 2 1 2 ...
## $ building_area : num 126.7 86.9 47 119.7 104 ...
## $ directions1 : chr "南" "南" "南" "北" ...
## $ directions2 : chr "北" "" "" "东" ...
## $ decoration : chr "精装" "精装" "简装" "精装" ...
## $ property_t_height: int 17 28 18 32 34 34 7 34 5 7 ...
## $ property_height : chr "中" "中" "低" "高" ...
## $ property_style : chr "塔楼" "板楼" "塔楼" "塔楼" ...
## $ followers : int 3 1 3 2 3 1 0 0 2 0 ...
## $ near_subway : chr "近地铁" NA "近地铁" "近地铁" ...
## $ if_2y : chr NA "房本满两年" NA "房本满两年" ...
## $ has_key : chr "随时看房" "随时看房" "随时看房" "随时看房" ...
## $ vr : chr NA "VR看装修" NA NA ...
```

可以看到：

- 直观结论1：数据中共有11个字符型变量，2个数值型变量，5个整数值型变量。
- 直观结论2：数据从房屋单价1771元到44656元，房屋面积从22.77平方米到588.66平方米，两项数据的极值较大，房屋类型丰富。

探索性分析

房屋区域的数值描述与图形

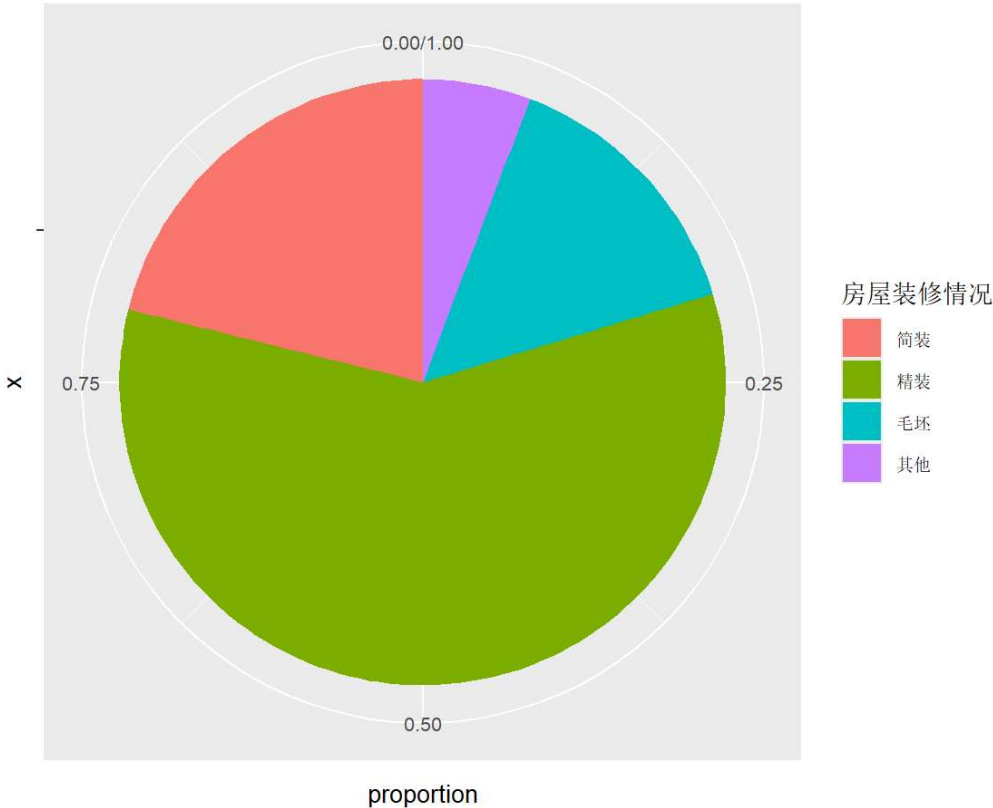


发现：

- 发现1：小区所处区域主要集中在白沙洲、盘龙城、四新、光谷东附近。推断这几个区域的城市主要服务设施较多，或是城市中心的商务区附近，所以房屋分布比较密集。
- 发现2：房屋的集中出现和区域从中心向外沿扩张可以看出城市的发展趋势，白沙洲、盘龙城、四新、光谷东等热门的区域可能有旧房改造和政府重点规划的新城或者开发区。

房屋装修情况的数值描述与图形

房屋装修情况饼状图

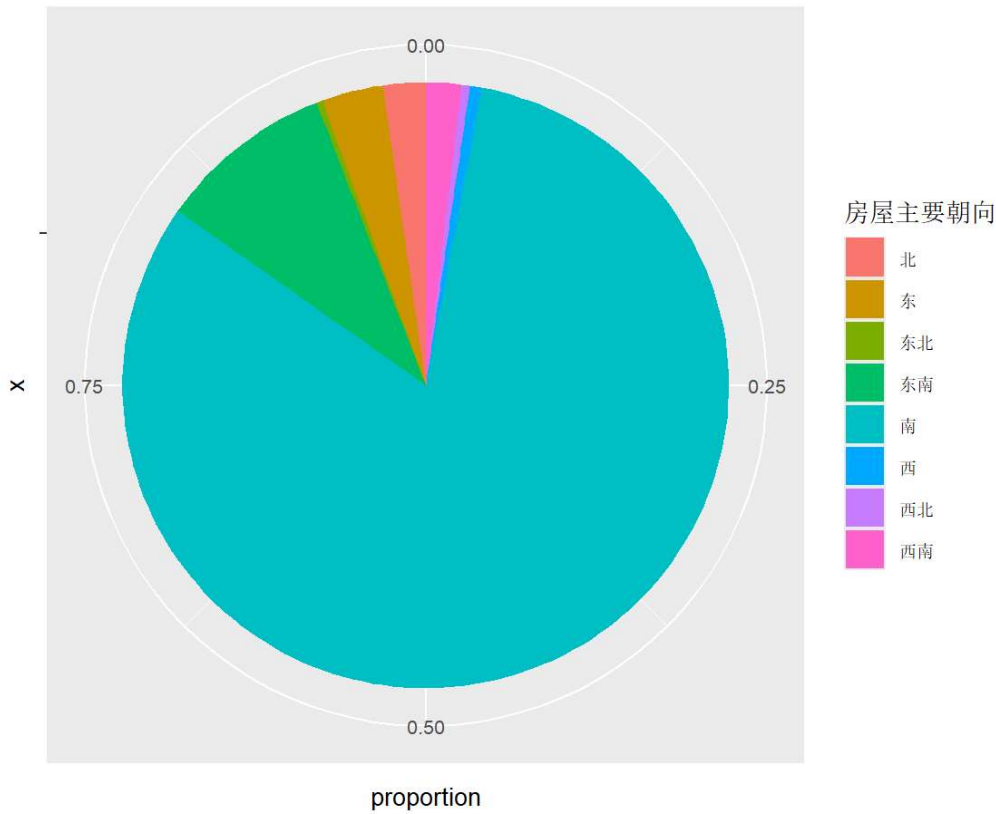


发现：

- 发现1：二手房交易市场大部分房子是已经装修过的，这一定程度上反映了城市实际居住人口多，城市经济发达。经济发达区域的购房者可能更注重时间成本和生活品质，房产投资者在这些区域也更倾向于将房子装修后再出租或出售，以获取更高的收益。

房屋主要朝向情况的数值描述与图形

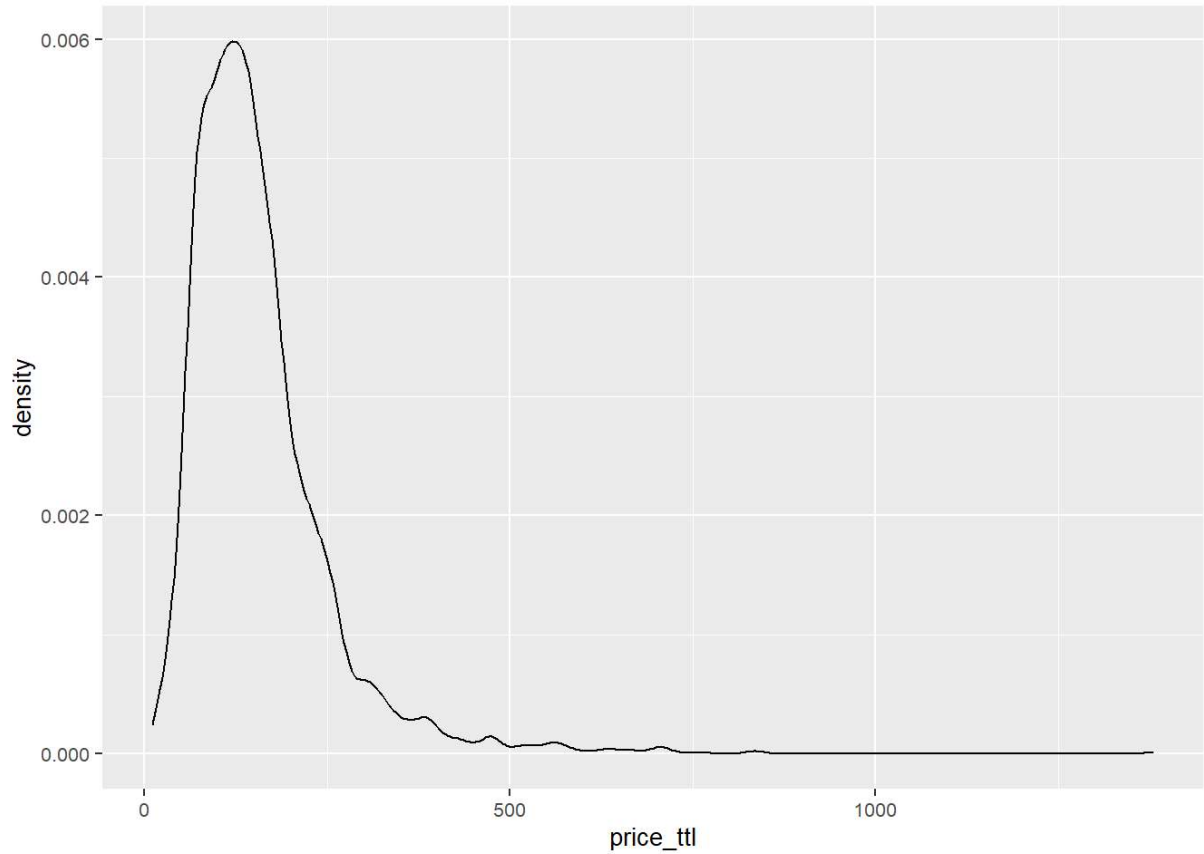
房屋主要朝向情况饼状图

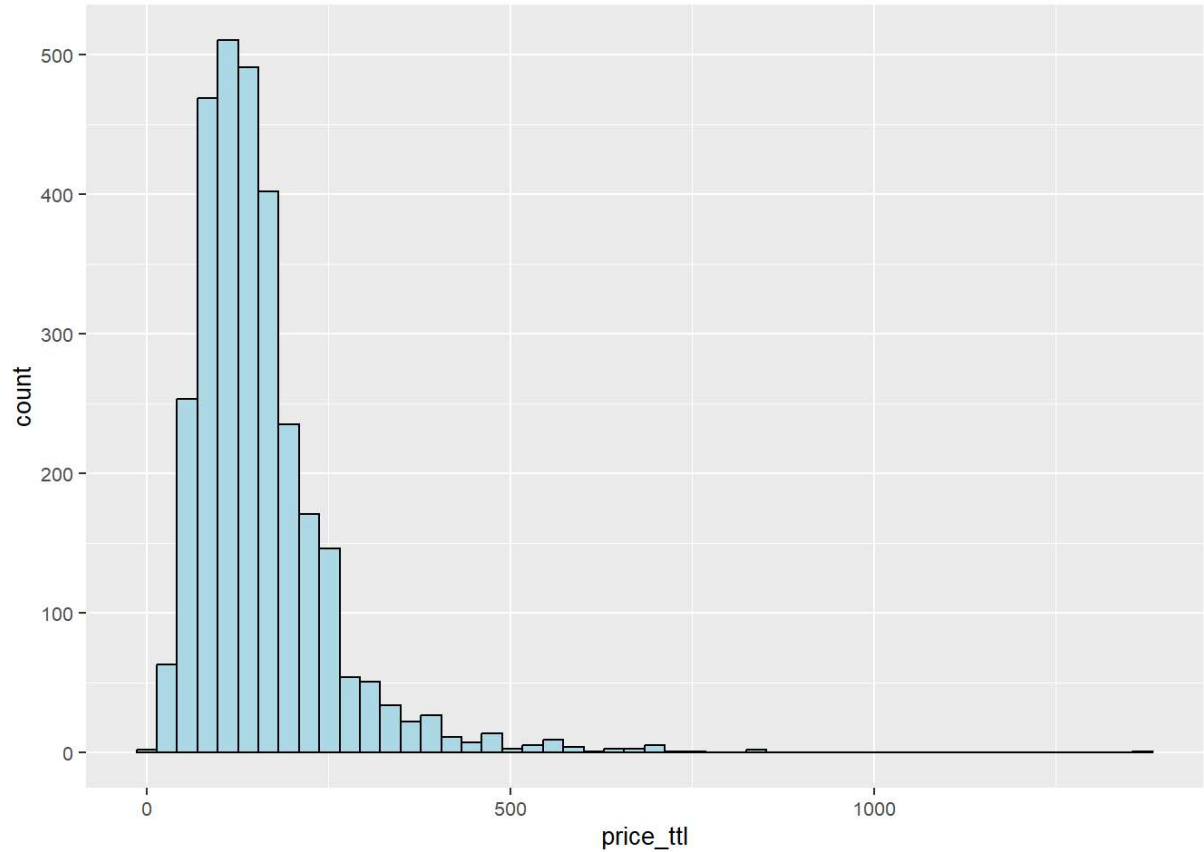


发现:

- 发现1：大部分房子的主要朝向是向西，其次是朝南方向。这也和城市人群在房屋中实际居住时间多半为下午和晚上有关，房屋主要朝西意味着在下午和傍晚能够获得较多的阳光照射，阳光能够深入室内，提供良好的自然采光和热量。

房屋总价的数值描述与图形





发现:

- 发现1：房屋总价的密度分布函数呈右偏趋势，从数据分布的角度看，右偏分布存在一个长长的“尾巴”，代表高价值房产。大部分房屋价格集中在较低到中等的价位区间，而有少数高价值房产拉高了整体的价格分布。
- 发现2：房屋总价的右偏分布在一定程度上反映了社会经济的不平等。高价值房产的拥有者往往具有更高的经济实力，而大部分居民只能购买价格较低的房产。这种差异可能与居民的收入分配不均有关，也反映了房地产市场在财富分配中的作用。

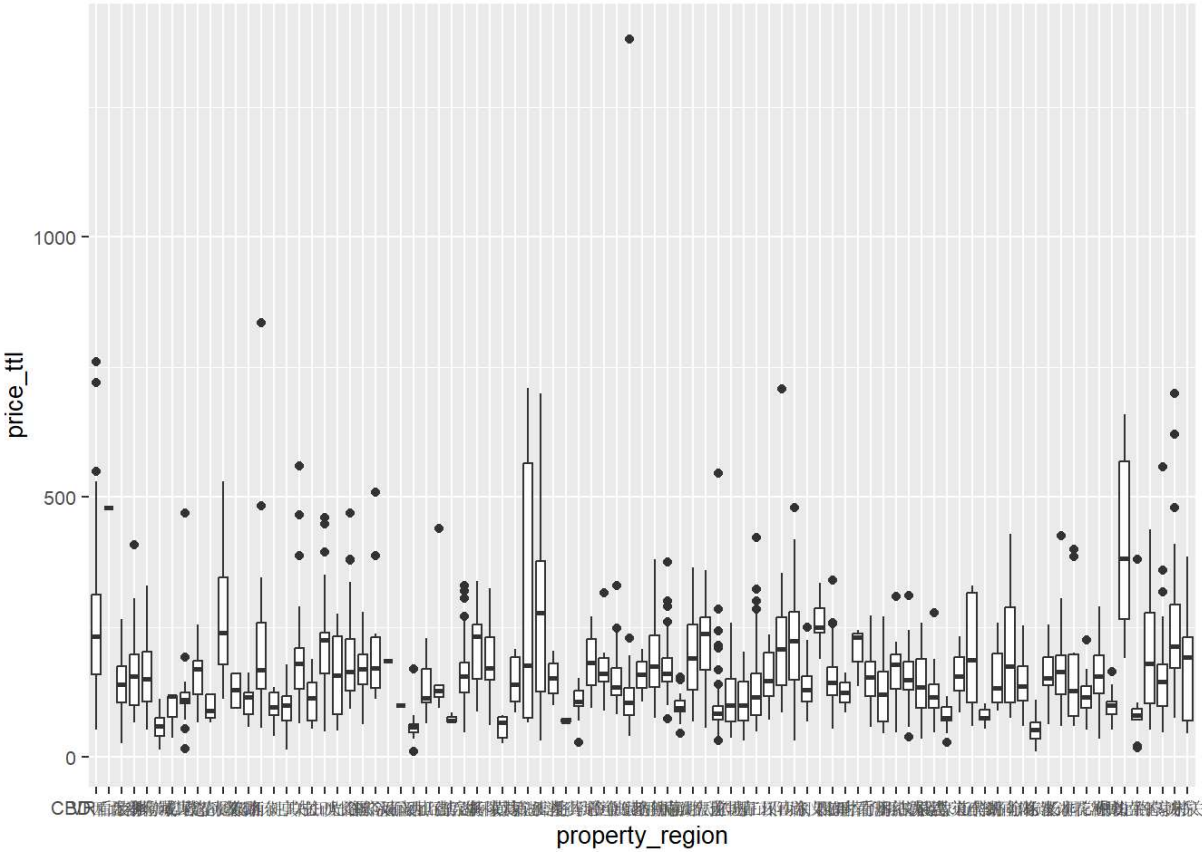
探索问题1：不同区域和建筑面积对房产总价格影响

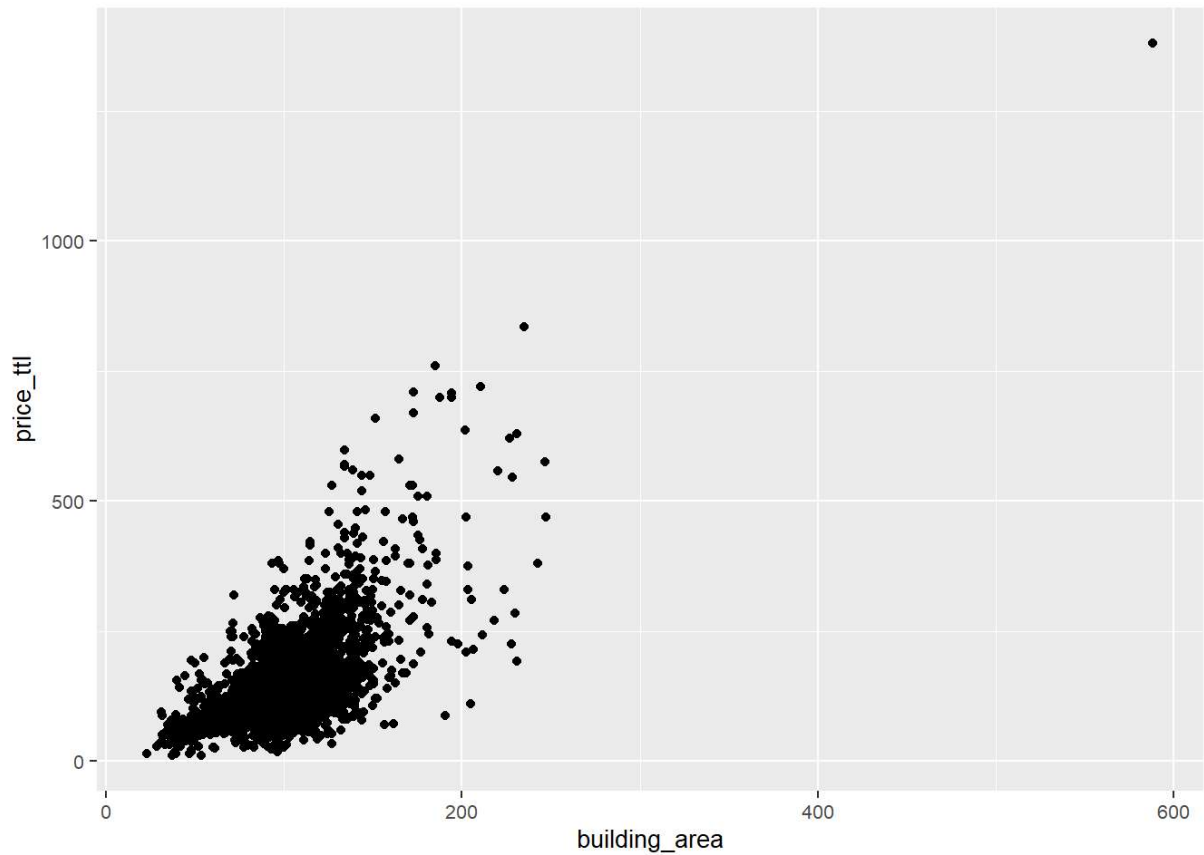
Data summary							
Name						lj1	
Number of rows						3000	
Number of columns						18	
<hr/>							
Column type frequency:							
character						11	
numeric						7	
<hr/>							
Group variables						None	
Variable type: character							
skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
property_name	0	1.00	2	17	0	1345	0
property_region	0	1.00	2	6	0	87	0
directions1	0	1.00	1	2	0	8	0
directions2	0	1.00	0	2	1672	9	0
decoration	0	1.00	2	2	0	4	0
property_height	60	0.98	1	1	0	3	0
property_style	0	1.00	2	4	0	5	0
near_subway	1441	0.52	3	5	0	5	0
if_2y	1736	0.42	5	5	0	1	0
has_key	458	0.85	3	5	0	9	0

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
vr	906	0.70	3	5	0	11	0

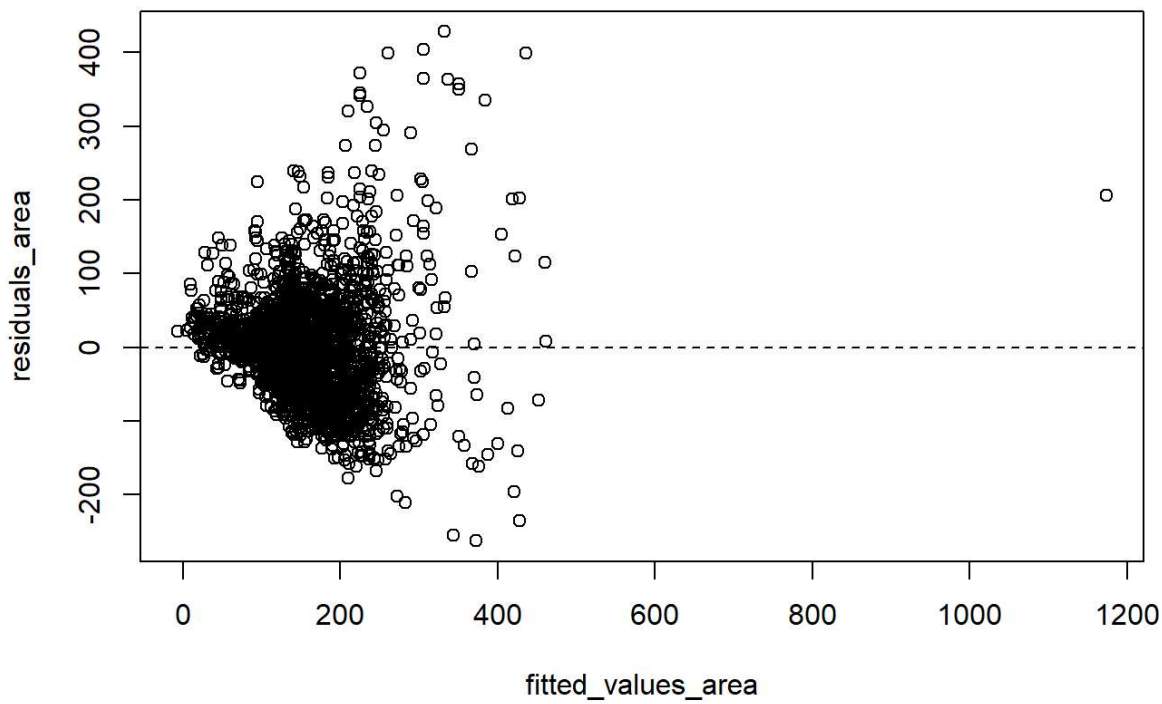
Variable type: numeric

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
price_ttl	0	1	155.86	95.55	10.60	95.00	137.00	188.00	1380.00	
price_sqm	0	1	15148.49	6323.18	1771.00	10799.25	14404.00	18211.00	44656.00	
bedrooms	0	1	2.69	0.73	1.00	2.00	3.00	3.00	7.00	
livingrooms	0	1	1.71	0.47	0.00	1.00	2.00	2.00	4.00	
building_area	0	1	100.87	30.38	22.77	84.92	95.54	117.68	588.66	
property_t_height	0	1	24.22	12.45	2.00	11.00	27.00	33.00	62.00	
followers	0	1	6.61	15.22	0.00	1.00	3.00	6.00	262.00	

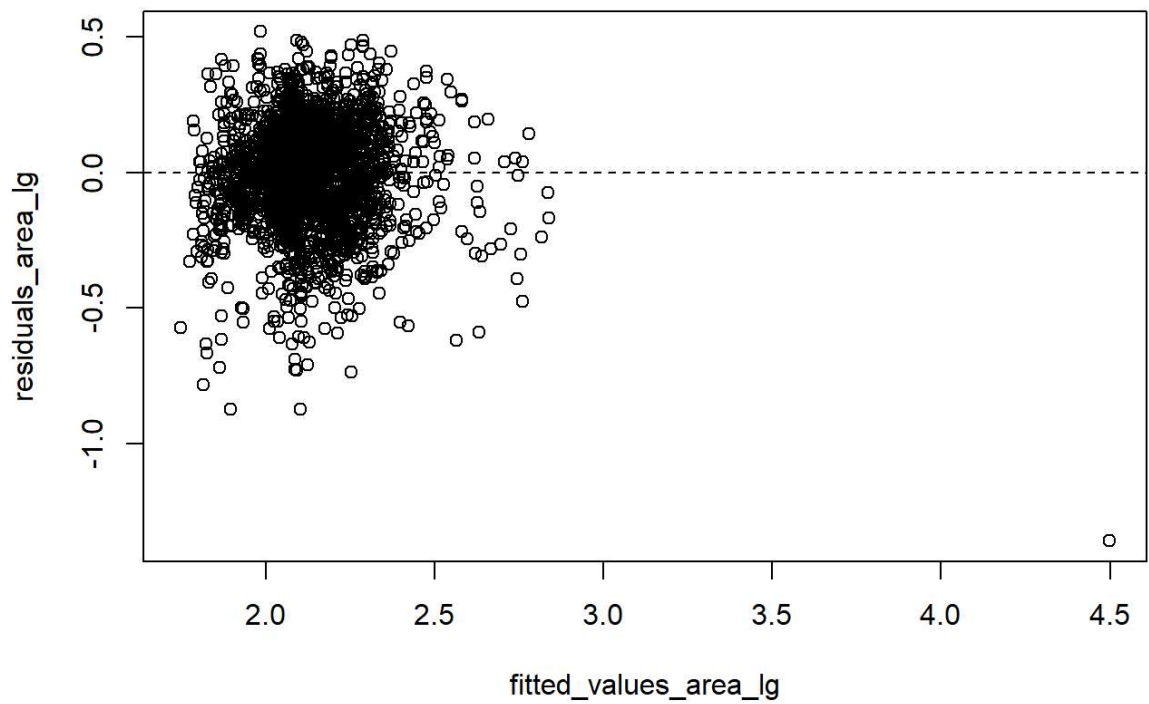




```
##
## Call:
## lm(formula = price_ttl ~ building_area, data = ljl)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -262.39  -42.13   -3.42   30.54   428.14
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -54.64516     4.52675  -12.07  <2e-16 ***
## building_area   2.08699     0.04297   48.57  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 71.49 on 2998 degrees of freedom
## Multiple R-squared:  0.4403, Adjusted R-squared:  0.4401
## F-statistic: 2359 on 1 and 2998 DF,  p-value: < 2.2e-16
```

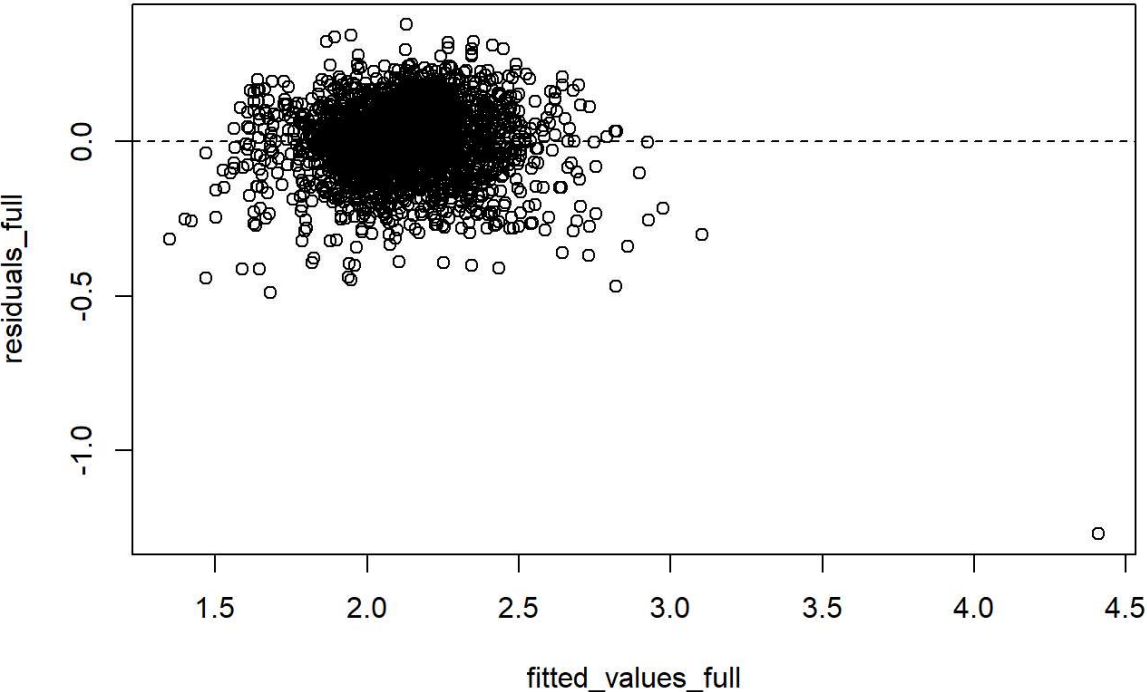


```
##
## Call:
## lm(formula = log10(price_ttl) ~ building_area, data = ljl)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.35821 -0.11242  0.01894  0.12423  0.51958
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   1.6373143   0.0118872   137.74  <2e-16 ***
## building_area  0.0048598   0.0001128    43.07  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1877 on 2998 degrees of freedom
## Multiple R-squared:  0.3822, Adjusted R-squared:  0.382
## F-statistic: 1855 on 1 and 2998 DF,  p-value: < 2.2e-16
```

```
##
## Call:
## lm(formula = log10(ljl$price_ttl) ~ ., data = as.data.frame(real_estate_matrix))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.26889 -0.07134  0.01062  0.08071  0.37654
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.773e+00  2.234e-02   79.367  < 2e-16 ***
## ` (Intercept)`              NA              NA              NA              NA
## property_regionVR看装修      2.831e-01  1.231e-01    2.300  0.021546 *
## property_region白沙洲     -1.543e-01  2.260e-02  -6.828  1.05e-11 ***
## property_region百步亭     -8.391e-02  2.611e-02  -3.214  0.001322 **
## property_region宝丰崇仁    -7.810e-02  3.307e-02  -2.362  0.018243 *
## property_region蔡甸城区    -5.474e-01  2.865e-02 -19.109  < 2e-16 ***
## property_region蔡甸其它    -4.553e-01  7.303e-02  -6.235  5.18e-10 ***
## property_region藏龙岛     -2.831e-01  2.902e-02  -9.757  < 2e-16 ***
## property_region常青花园    -8.107e-02  2.886e-02  -2.809  0.005005 **
## property_region常青路     -2.366e-01  4.358e-02  -5.428  6.16e-08 ***
## property_region楚河汉街     1.752e-02  3.746e-02   0.468  0.640019
## property_region大智路     -1.907e-01  5.027e-02  -3.793  0.000152 ***
## property_region堤角       -2.056e-01  2.869e-02  -7.168  9.59e-13 ***
## property_region东湖东亭    -2.152e-02  2.846e-02  -0.756  0.449636
## property_region东西湖其它  -2.934e-01  5.028e-02  -5.834  6.00e-09 ***
## property_region沌口       -2.983e-01  2.638e-02 -11.308  < 2e-16 ***
## property_region二七       -1.461e-02  2.768e-02  -0.528  0.597599
## property_region古田       -1.852e-01  2.803e-02  -6.608  4.61e-11 ***
## property_region关山大道     2.705e-02  3.180e-02   0.851  0.395030
## property_region关西长职    -7.689e-02  3.266e-02  -2.354  0.018643 *
## property_region光谷东     -8.140e-02  2.351e-02  -3.462  0.000544 ***
## property_region光谷广场    -8.622e-02  3.180e-02  -2.711  0.006746 **
## property_region光谷南     -6.789e-02  3.110e-02  -2.183  0.029099 *
## property_region国际百纳    -9.576e-02  1.231e-01  -0.778  0.436746
## property_region汉口北     -3.994e-01  1.231e-01  -3.245  0.001190 **
## property_region汉南其它    -5.721e-01  3.664e-02 -15.612  < 2e-16 ***
## property_region汉正街     -1.470e-01  4.356e-02  -3.375  0.000748 ***
## property_region洪山其它    -9.910e-02  5.365e-02  -1.847  0.064847 .
## property_region后官湖     -4.329e-01  7.303e-02  -5.928  3.43e-09 ***
## property_region后湖       -1.373e-01  2.434e-02  -5.639  1.87e-08 ***
## property_region虎泉杨家湾  6.812e-02  3.355e-02   2.031  0.042394 *
## property_region华科大     -4.222e-02  3.145e-02  -1.343  0.179479
## property_region黄陂其它    -5.401e-01  5.366e-02 -10.066  < 2e-16 ***
## property_region黄家湖     -2.162e-01  4.353e-02  -4.967  7.18e-07 ***
## property_region黄埔永清     8.732e-03  3.260e-02   0.268  0.788802
## property_region积玉桥     -2.893e-02  2.559e-02  -1.130  0.258453
## property_region集贤       -1.105e-01  4.354e-02  -2.538  0.011214 *
## property_region江夏其它    -4.616e-01  5.364e-02  -8.606  < 2e-16 ***
## property_region将军路     -2.455e-01  3.463e-02  -7.091  1.66e-12 ***
## property_region街道口     2.626e-02  8.829e-02   0.297  0.766175
## property_region街道口     -4.504e-02  3.946e-02  -1.142  0.253708
## property_region金融港     -8.521e-02  2.659e-02  -3.204  0.001370 **
## property_region金银湖     -3.046e-01  2.394e-02 -12.724  < 2e-16 ***
## property_region近地铁     -1.491e-01  7.303e-02  -2.041  0.041322 *
## property_region老南湖     -1.584e-02  2.638e-02  -0.601  0.548202
## property_region珞狮南路    -9.050e-02  2.532e-02  -3.574  0.000357 ***
## property_region庙山       -3.158e-01  2.970e-02 -10.633  < 2e-16 ***
## property_region民族大道    -1.116e-02  2.578e-02  -0.433  0.665106
## property_region南湖沃尔玛  6.455e-02  2.951e-02   2.188  0.028761 *
## property_region盘龙城     -3.794e-01  2.320e-02 -16.351  < 2e-16 ***
## property_region七里庙     -2.261e-01  2.908e-02  -7.775  1.04e-14 ***
## property_region前进江汉    -1.419e-01  3.480e-02  -4.078  4.67e-05 ***
## property_region青山       -1.529e-01  2.446e-02  -6.250  4.70e-10 ***
## property_region三环南     -8.479e-02  3.524e-02  -2.406  0.016202 *
## property_region三阳路      7.385e-02  3.667e-02   2.014  0.044145 *
## property_region沙湖       -3.626e-03  2.969e-02  -0.122  0.902825
## property_region首义       -1.375e-01  3.222e-02  -4.266  2.05e-05 ***
## property_region水果湖     1.620e-01  4.540e-02   3.568  0.000365 ***
## property_region四新       -1.181e-01  2.344e-02  -5.039  4.97e-07 ***
## property_region随时看     -7.689e-02  8.831e-02  -0.871  0.384016
## property_region随时看房   -1.044e-02  7.303e-02  -0.143  0.886334
## property_region塔子湖     -8.280e-02  2.513e-02  -3.294  0.000999 ***
## property_region台北香港路 -1.716e-01  3.224e-02  -5.324  1.09e-07 ***
## property_region唐家墩     -6.601e-02  3.526e-02  -1.872  0.061261 .
## property_region团结大道    -9.673e-02  2.549e-02  -3.794  0.000151 ***
## property_region王家湾     -1.614e-01  2.473e-02  -6.527  7.87e-11 ***
## property_region文化大道    -2.303e-01  2.540e-02  -9.067  < 2e-16 ***
## property_region吴家山     -3.653e-01  3.356e-02 -10.887  < 2e-16 ***
## property_region武昌火车站 -1.053e-01  3.352e-02  -3.141  0.001698 **
## property_region武广万松园  5.198e-04  4.541e-02   0.011  0.990868
## property_region武湖       -3.985e-01  3.747e-02 -10.635  < 2e-16 ***
## property_region新华路万达 -6.685e-02  3.668e-02  -1.823  0.068463 .
```

```
## property_region新南湖      -1.410e-02  3.051e-02  -0.462  0.644092
## property_region徐东      -1.214e-01  2.715e-02  -4.469  8.15e-06 ***
## property_region阳逻      -6.078e-01  2.541e-02 -23.926  < 2e-16 ***
## property_region杨汊湖    -1.157e-01  2.947e-02  -3.926  8.83e-05 ***
## property_region杨园      -7.490e-02  3.305e-02  -2.266  0.023502 *
## property_region育才花桥  -7.545e-02  3.527e-02  -2.139  0.032518 *
## property_region长丰常码头 -2.142e-01  2.905e-02  -7.375  2.14e-13 ***
## property_region长港路    -7.362e-02  2.783e-02  -2.646  0.008198 **
## property_region纸坊      -3.297e-01  2.995e-02 -11.009  < 2e-16 ***
## property_region中北路     1.752e-01  3.521e-02   4.975  6.89e-07 ***
## property_region中法生态城 -5.108e-01  4.061e-02 -12.578  < 2e-16 ***
## property_region中南丁字桥 -2.292e-03  2.671e-02  -0.086  0.931613
## property_region钟家村     -1.247e-01  2.549e-02  -4.893  1.04e-06 ***
## property_region卓刀泉     -9.527e-03  3.020e-02  -0.315  0.752466
## property_region宗关       -9.602e-02  2.926e-02  -3.281  0.001046 **
## building_area            4.995e-03  7.574e-05  65.947  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1214 on 2912 degrees of freedom
## Multiple R-squared:  0.7491, Adjusted R-squared:  0.7416
## F-statistic: 99.95 on 87 and 2912 DF,  p-value: < 2.2e-16
```



发现:

- 发现1: 建筑面积与房产总价的对数呈正相关, 这意味着在其他条件相同的情况下, 建筑面积越大, 房产总价越高。从经济角度看, 房产面积的增加带来了价值的超比例增长 (因为是对数关系)。例如, 可能每增加一定比例的建筑面积, 房产总价会以一个相对稳定的比例增长。这可能是由于较大的空间能满足更多的功能需求, 如容纳更多的居住人口、提供更宽敞的商业经营场所等, 从而使得其价值提升。
- 发现2: 暗示区域的发展潜力不同。具有良好经济发展前景、城市规划重点关注的区域, 房产总价往往更高。例如, 新兴的商务区或者正在进行大规模城市更新的区域, 房产价值可能随着区域的发展而不断提升, 从而体现出与房产总价对数的正相关关系。
- 发现3: 这种显著的正相关关系反映了房地产市场的定价规律。市场在定价过程中会充分考虑房产的物理属性 (建筑面积) 和地理位置属性 (区域)。消费者和投资者在购买房产时也会基于这些因素来衡量房产的价值, 并且这种价值衡量在宏观市场层面上表现为对数形式的正相关关系, 为房地产市场的价格评估和预测提供了一定的参考依据。

探索问题2: 房屋在所在楼栋所处位置、装修情况对房产单价的影响

```
##           Df    Sum Sq   Mean Sq F value Pr(>F)
## property_height  2  7.493e+07  3.747e+07   0.966  0.381
## decoration      3  3.735e+09  1.245e+09  32.099 <2e-16 ***
## Residuals      2934  1.138e+11  3.879e+07
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 因为不存在, 60个观察量被删除了
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = price_sqm ~ property_height + decoration, data = ljl)
##
## $decoration
##          diff          lwr          upr          p adj
## 精装-简装 2087.2259 1339.009 2835.4428 0.0000000
## 毛坯-简装 -232.6641 -1243.368 778.0398 0.9346096
## 其他-简装 -708.3326 -2090.366 673.7011 0.5518070
## 毛坯-精装 -2319.8900 -3190.894 -1448.8858 0.0000000
## 其他-精装 -2795.5585 -4078.969 -1512.1477 0.0000001
## 其他-毛坯 -475.6685 -1927.852 976.5150 0.8344000
```

发现:

- 发现1: 装修情况对房产单价有显著影响, 且精装和其他类型存在显著差异, 这表明装修是房产价值的重要驱动因素。精装修的房产能够显著提升房产单价, 可能是因为精装修节省了购房者的时间和精力, 并且在装修质量、风格统一等方面有一定优势, 使得其在市场上更具竞争力, 能够获得更高的价格认可。消费者可能更愿意为精装修的房产支付较高的单价, 因为精装修的房产在入住时可以直接使用, 避免了装修过程中的诸多麻烦, 如寻找装修公司、监督施工、选择装修材料等。这也暗示房地产开发商或二手房卖家可以通过提供高质量的精装修来增加房产的附加值。
- 发现2: 房屋楼层高度对房产单价影响不显著, 说明在市场定价过程中, 楼层因素并没有像装修情况那样起到关键作用。这可能是因为在该研究的 market 环境中, 楼层所带来的优势 (如视野好、噪音小等) 和劣势 (如使用电梯不便、楼层高水压不足等) 相互抵消, 或者这些因素在消费者购买决策中的权重较低, 导致楼层高度没有对房产单价产生显著的分化作用。

探索问题3: 房屋的便利性因素 (如靠近地铁、产证满 2 年、随时看房、支持 VR看房) 是否会增加其在网站上的关注度

```
##
## Call:
## glm(formula = followers ~ near_subway + if_2y + has_key + vr,
##      family = poisson, data = ljl)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      1.11507    0.02787  40.005 < 2e-16 ***
## near_subwayVR看装修 1.72742    0.16420  10.520 < 2e-16 ***
## near_subway近地看 0.96129    0.26784   3.589 0.000332 ***
## near_subway近地铁 0.29338    0.01457  20.131 < 2e-16 ***
## near_subway珞狮南 1.00943    0.23611   4.275 1.91e-05 ***
## near_subway太子湖1号 -1.05935    0.70725  -1.498 0.134174
## if_2y房本满两年 0.12844    0.01445   8.889 < 2e-16 ***
## has_keyVR看装修 0.28561    0.22496   1.270 0.204224
## has_key藏龙岛 -0.79003    0.70756  -1.117 0.264184
## has_key关山大道 -12.62082   172.65451  -0.073 0.941728
## has_key近地铁 1.10786    0.11807   9.383 < 2e-16 ***
## has_key仁厚社区 -12.71103   172.65451  -0.074 0.941312
## has_key世纪花园 -0.49665    0.70760  -0.702 0.482757
## has_key随时看房 0.56270    0.02466  22.817 < 2e-16 ***
## has_key随李铁 2.41373    0.13619  17.724 < 2e-16 ***
## has_key王家湾 -12.62082   172.65451  -0.073 0.941728
## vrVR\小区\新城\8c -1.40645    0.70732  -1.988 0.046766 *
## vr随时看\88 3.05843    0.09529  32.095 < 2e-16 ***
## vr保利拉\8f 0.09763    0.33379   0.292 0.769909
## vr育才花\8a1 -0.87254    0.57757  -1.511 0.130861
## vrVR看装修 0.07473    0.01594   4.689 2.75e-06 ***
## vr泓悦府 -0.57916    0.57762  -1.003 0.316022
## vr江景湾 -0.49016    0.44755  -1.095 0.273431
## vr近地铁 0.13970    0.37841   0.369 0.712000
## vr珞狮南路 -1.67777    1.00015  -1.678 0.093443 .
## vr随时看房 -12.98035   172.65451  -0.075 0.940071
## vr塔子湖 -0.98462    0.70733  -1.392 0.163911
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
## Null deviance: 41889  on 2999  degrees of freedom
## Residual deviance: 40057  on 2973  degrees of freedom
## AIC: 48078
##
## Number of Fisher Scoring iterations: 9
```

发现:

- 发现1: 靠近地铁这一因素有显著影响, 说明对于关注看房的人群来说, 交通便利性是他们考虑的重要因素。在购房或者租房决策中, 靠近地铁意味着更便捷的通勤条件, 减少交通成本和时间成本。。
- 发现2: 靠近地铁这一因素有显著影响, 说明对于关注看房的人群来说, 交通便利性是他们考虑的重要因素。在购房或者租房决策中, 靠近地铁意味着更便捷的通勤条件, 减少交通成本和时间成本。这也反映出大部分看房者可能是上班族或者对交通便捷性有较高要求的人群, 房产周边的交通设施对他们的吸引力较大。

- 发现3：随时看房（中介有钥匙）和支持VR看房这两个因素的显著性，体现了看房者对于看房便利性的重视。随时看房为他们提供了更灵活的看房时间安排，不需要等待房东或者其他复杂的协调过程。VR看房则适应了现代科技发展和快节奏生活的需求，能够让看房者在初步筛选阶段通过虚拟现实技术对房屋有一个直观的了解，节省时间和精力。

发现总结

用1-3段话总结你的发现。

通过对武汉链家二手房的数据分析可以看出： 1.在购房或租房决策中，靠近地铁对看房人群有显著影响，交通便利性至关重要，能减少通勤成本和时间，反映出多数看房者可能是上班族或对交通便捷要求高的人。随时看房和VR看房的显著性体现看房者对看房便利性的重视，前者提供灵活安排，后者节省时间精力。 2.装修情况对房产单价有显著影响，精装修能提升单价，因其节省购房者时间精力、质量风格有优势，开发商和卖家可借此增加附加值，而楼层高度对单价影响不显著。 3.建筑面积与房产总价对数呈正相关，面积增加带来价值超比例增长，同时区域发展潜力影响房产总价，发展前景好、受规划关注区域总价更高，这些正相关关系反映房地产市场定价规律，为价格评估和预测提供参考。