

基于主成分分析的中国热门城市房地产的投资潜力评价

一、研究房地产的投资潜力的目的

1998 年,我国结束福利分房并全面推行住房商品化,2003 年国务院出台的“18 号文件”明确将房地产定位为国民经济支柱产业,从此,房地产行业进入了快速发展阶段,房地产行业的回报率水涨船高。本文收集了上一年 20 个中国热门城市的部分宏观经济数据和房地产市场数据,利用主成分分析对这些城市的房地产投资潜力作综合评价,力图找到未来几年最具有投资潜力的城市,这对希望投资房地产行业的企业或个人有一定的参考价值。

二、房地产投资潜力的评价指标体系

本文综合考虑影响城市房地产市场的各个方面的因素,遵循数据的可获得性、全面性、可比较性等原则,选择了反映一个城市经济发展水平,人口因素,房地产市场景气程度大维度的三大维度 8 个指标,构建城市房地产投资潜力评价指标体系,同时收集了 20 个城市的数据,对这 20 个城市的房地产投资潜力进行综合评价。

表 1 房地产投资潜力的评价指标体系

评价内容	变量名	单位	取值范围
经济发展水平	GDP	亿元	333~10297
	人均可支配收入	元	13350~25320
	城市化水平	无	52~88
人口因素	人均使用面积	平方米	16~45
	户籍人口数量	万人	92~1368
房地产市场	商品房销售均价	元/平方米	4134~9230
景气程度	商品房销售面积	万平方米	91~3025
	商品房施工面积	万平方米	300~10939

三、相关指标的描述统计分析

描述统计分析先对各个指标的样本分布进行分析,绘制分布直方图,再通过相关图观察各个指标之间的相关性。

(一) 指标的分布情况

由图 1 各个指标的分布可以看出,各个指标均呈右偏分布,每个指标下都有一到两个城市特别突出。其中上海排名第一的次数有 5 次,舟山排名最末的次数有 3 次,上海在人均使用面积这个指标下排名最末,预示这个指标可能为负指标。由各个指标的变异系数可以看出,各个城市在 GDP、商品房销售面积、商品房施

工面积这三个指标下离散程度大，而在人均可支配收入、城市化水平这两个指标下离散程度小。

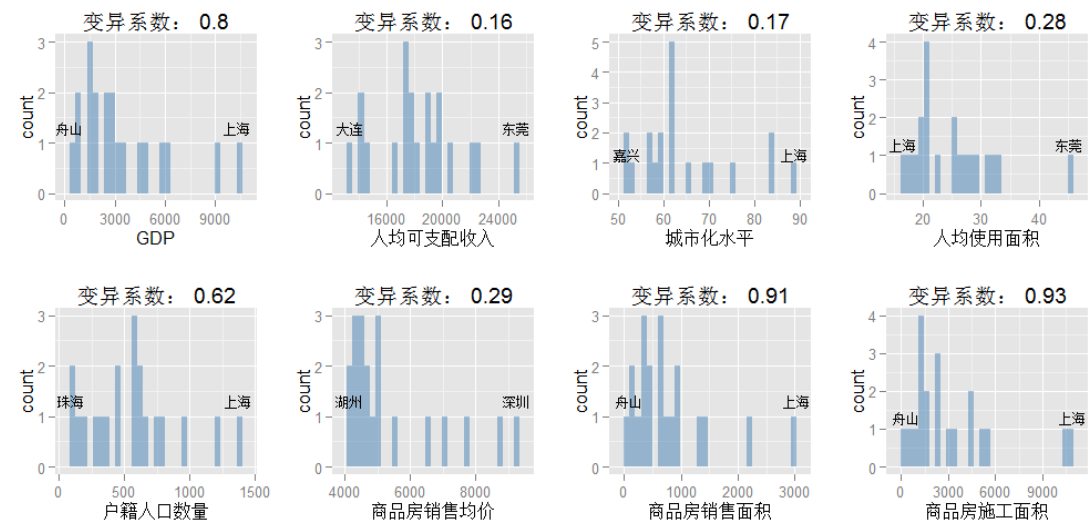


图 1 各个指标下的样本分布

(二) 指标相关性情况

如相关图图 2 所示，指标间的相关系数为正，则为蓝色，为负则为红色，蓝色越深表明指标间正相关性越强。从图 2 可以看出，人均使用面积与其他指标均为负相关关系，说明这个指标是负指标，其他均为正指标。商品房销售面积、施工面积、GDP 三者之间为强相关关系，它们两两的的相关系数均达到 0.9 以上。由于最后两个指标都是反映房地产市场的景气程度，与其他指标的相关性十分相像，如果都放进后面的主成分分析中会造成房地产市场景气程度这方面的信息重叠，影响权重，因此考虑删除其中一个指标。

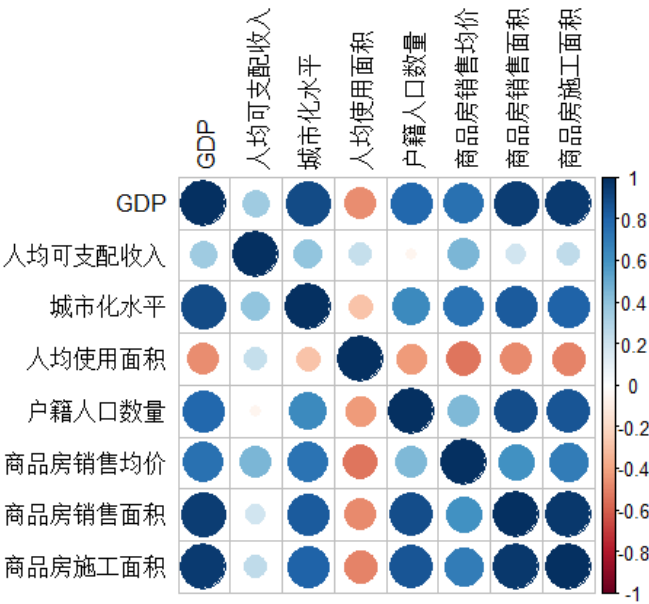


图 2 各个指标之间的相关图

四、 房地产投资潜力的主成分分析

主成分分析部分先对数据进行处理,使之适用于主成分分析和综合评价过程,再做通过主成分系数、得分等确定综合评价的方式。

(一) 数据处理

1. 指标正向化

由描述统计的结果知,人均使用面积是负指标,需要通过取倒数的方法正向化,因此改为 1/人均使用面积。

2. 无量纲化

做主成分分析时常用的无量纲化方法为标准化,但这会把各个样本在不同指标间的变异抹杀,影响主成分的方差贡献率,影响综合评价过程,因此此处用均值化的方法。

3. 指标筛选

由描述统计的结果可知,商品房销售面积与商品房施工面积的相关系数均达到 0.9 以上,这两个指标都是反映房地产市场的景气程度,与其他指标的相关性十分相像,如果都放进后面的主成分分析中会造成房地产市场景气程度这方面的信息重叠。考虑到本文的研究目的是寻找未来几年房地产市场具有良好投资潜力的城市,商品房施工面积预示未来房地产市场的商品房销售面积,因此删去商品房销售面积,保留商品房施工面积。

(二) 主成分分析

用均值化后的数据的协方差求其特征根,即为各个主成分的方差,反映它们的重要程度。由表 2 可知,到第二个主成分时,累计方差贡献率已达到 95.55%,图 3 碎石图的拐点在第三个主成分,这都预示着只需要取前两个主成分即可。

表 2 各个主成分的重要性

	Comp.1	Comp.2	Comp.3	Comp.4
标准差	4225.10	2681.56	945.66	501.82
方差贡献率	68.11%	27.44%	3.41%	0.96%
累计方差贡献率	68.11%	95.55%	98.96%	99.93%

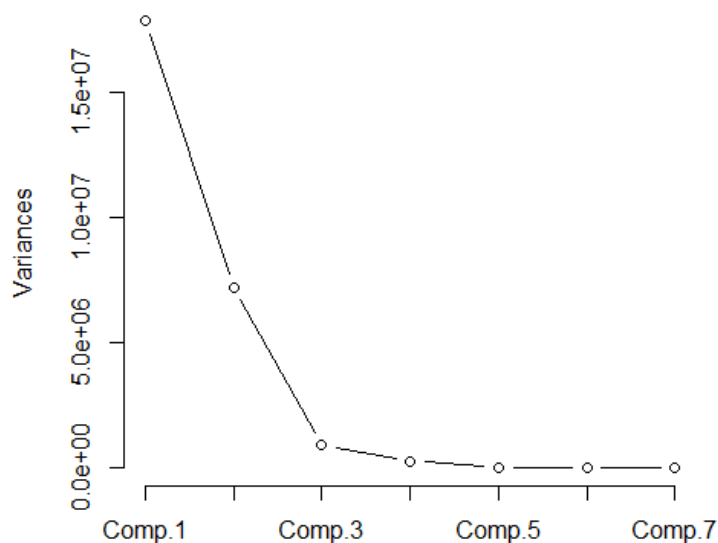


图 3 主成分碎石图

由表 3 可知，第一个主成分系数全为负，因此计算主成分得分时需要全部变为正数。GDP，商品房施工面积在第一主成分的系数达到 0.6 以上，因此第一主成分体现城市总体的宏观经济和房地产市场的景气程度，人均可支配收入在第二主成分上的系数达到 0.9 以上，因此第二主成分体现城市人民的购买力。

表 3 前两个主成分系数表

指标	Comp.1	Comp.2
GDP	-0.6	-0.22
人均可支配收入	-0.39	0.91
城市化水平	0	0
人均使用面积	0	0
户籍人口数量	-0.06	-0.07
商品房销售均价	-0.29	0.02
商品房施工面积	-0.64	-0.35

由图 4 左图所示，由于城市化水平、人均使用面积、户籍人口数量在第一、第二主成分上为 0 或接近于 0，因此在图中的位置非常接近，GDP、商品房施工面积位置近，说明这两个变量较为相似。

由图 4 右图所示，第二主成分上系数最大值 0.91 为人均可支配收入，而最小值为对应的指标为商品房施工面积，这是一个负数。受此影响，北京，上海，广州在第二主成分得分上的排名较后，分别为 13,12,16。（北京，上海，广州在原数据集上人均可支配收入排名为 4,5,6）这就说明使用第一、第二主成分合成的排名会存在一定问题。

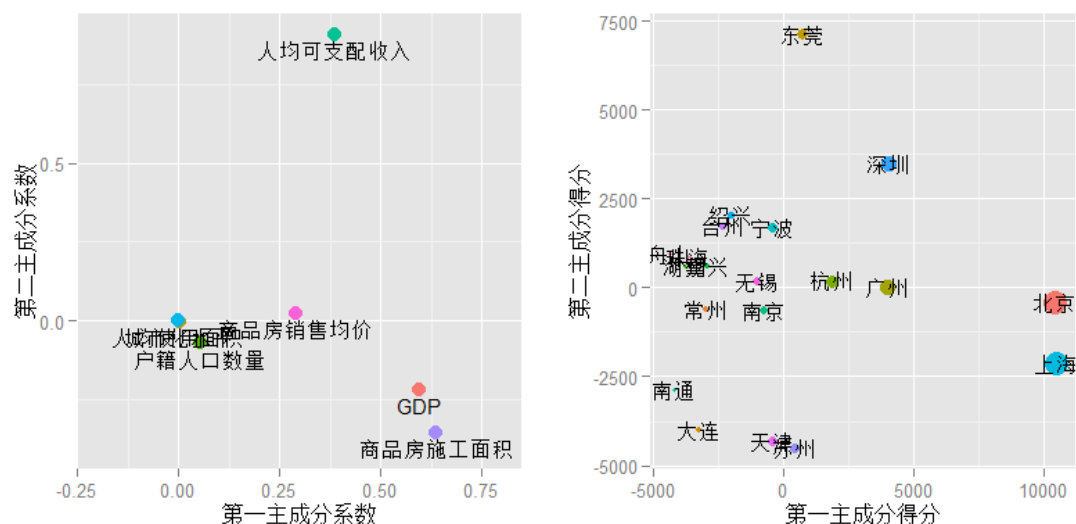


图 4 前两个主成分系数、得分图

五、 房地产投资潜力的综合评价结果与结论

由之前分析可知，使用第一、第二主成分合成的排名会使评价结果失真，且指标体系的三个方面均有在第一主成分上系数不为 0 的指标，因此用第一主成分对城市的房地产投资潜力排名更为合理。排名前十的城市及其第一主成分得分如表 4 所示。

表 4 房地产投资潜力综合评价前十名城市

排名	城市	第一主成分得分
1	上海	10483.29
2	北京	10406.38
3	深圳	4046.93
4	广州	3970.97
5	杭州	1848.93
6	东莞	734.45
7	苏州	417.37
8	天津	-403.06
9	宁波	-423.68
10	南京	-745.45

这份排名显示，北上广深一直都是全国最具房地产投资潜力的四大城市，与杭州、东莞、苏州拉开较大差距，但这三个城市也远远甩开剩余的 13 个城市，它们的潜力亦不容小觑。这在原数据集中，北京，上海在所有指标上的单独排名均在前 5，广州、深圳均在前 8，杭州均在前 10，东莞在人均可支配收入排名第一，在其他指标上的排名均有不俗表现。因此这份房地产投资潜力排名在较大程度反映了原数据的整体情况，对希望投资房地产行业的企业或个人有一定的参考价值。