

摘要

“拍照赚钱”APP 是基于移动互联网的自助式劳务众包平台，使得企业可利用大众力量，低成本、高效率地完成各种商品检查与信息搜集的任务。本文通过建立数学模型，就 APP 中的任务定价问题进行分析，给出最优的任务定价方案。

针对问题一, 针对问题的这一段总结。利用什么软件 + 建立什么模型 + 算法方法 + 求解的特点 + 主要的结果 + 评价及其推广。注意黑体字描黑重点关键词。

针对问题二, 对项目的任务定价规律进行**定性**与**定量**研究。利用 Matlab 的 cftool 工具箱绘制出任务的经纬度坐标与定价数据的**三维拟合图**, 观察到任务分布密集的地区任务定价较低。对任务的位置数据进行空间离散化处理和 **K-Means** 分析, 将任务分布的区域等划分为若干网格区域, 定义影响任务定价的四个因子, 即网格内任务数量、会员人数、会员平均完成能力、任务与中心点的距离。运用灰色关联矩阵定量分析四个影响因子与定价的相关度, 分别为 0.9710,0.9671,0.9633,0.9390。得出所定义的指标对定价相关性很高, 能较好描述定价规律。最后通过比较未完成任务与已完成任务的相关度矩阵得出距离对任务的完成的影响是最显著的

针对问题三,

针对问题四,

本文中所提到的模型优点主要有两点: 一、在与污染源头距离较短时预测抗噪能力较强; 二、利用更高定位精度和鲁棒性的直线解析法, 溯源追踪能力较强。

关键词: 对流扩散方程 直线解析法 溯源算法 拉普拉斯变换

目录

一、 问题重述	4
1.1 问题背景	4
1.2 问题提出	4
二、 模型假设	4
三、 符号说明	5
四、 问题一模型的建立与求解	5
4.1 问题的描述与分析	5
4.2 预备工作	6
4.2.1 评价指标体系构建	6
4.2.2 同向化处理	7
4.2.3 标准化处理	7
4.2.4 KMO 与 Bartlett 检验	7
4.3 模型的建立与求解	8
4.3.1 公因子的确定	8
4.3.2 因子命名与载荷矩阵计算	10
4.3.3 求得因子得分和综合绩效得分	12
4.4 结果分析	13
五、 问题二模型的建立与求解	13
5.1 问题的描述与分析	13
5.2 模型的建立与求解	14
5.2.1 人才及其吸引力影响因素的分类	14
5.2.2 分类型人才吸引力分析	14
5.3 结果分析	15
六、 问题三的求解	15
6.1 写给武汉市人力资源部门的建议报告	15
七、 模型的评价	17
7.1 模型的优点	17
7.2 模型的缺点	17
7.3 模型的改进与展望	17

附录 A 代码	18
A.1 数据预处理—python 源代码	18
A.2 数据可视化—python 源代码	18
A.3 相关系数矩阵	22
A.4 初始因子载荷矩阵	22
A.5 旋转后的因子载荷矩阵	24
A.6 成分得分系数矩阵	26

一、问题重述

1.1 问题背景

现代管理中,彼得德鲁克提出了”人力资源”的概念并认为”人力资源是与其他资源不一样的特殊资源”,人力资源的其中一个特质是其具有自主流动性。根据勒温的场论,人才会根据自身需求与环境的适应性而做出离开或者留在某个环境的决定。人才是一个城市保持竞争活力和创新力的关键,在世界各国和全国各地积极推出人才吸引的政策背景下,一个地区的人才吸引力的核心内涵就是这个地区所具有的可以影响人才选择的能力。

全面、科学、系统地评价一个城市的人才吸引力是制定人才吸引计划的前提,城市人才吸引评价指标的选取受多方面影响,关键是要符合人才的理想,满足人才的需求和愿望。按照重要程度,人才的需求可以分为发展前景,收入和环境。发展前景是首要关心的因素,收入是人才流动的另一关键因素,环境因素包括治安、交通、污染、教育、医疗、购物等,也都是会考虑的因素。

1.2 问题提出

围绕城市地区人才吸引力水平,以城市多指标评价体系为依据,依次提出以下问题:

- (1) 根据数学模型及收集的数据,定量地分析武汉市的人才吸引力水平,并就武汉市的人才政策对人才吸引力水平的影响作出定量评价。
- (2) 结合人才类别,针对不同类型的人才,深入分析比较武汉市与其他同类城市在人才吸引力上的优势与不足,并给出有效提升人才吸引力的可行方案。
- (3) 结合模型结果及分析,给武汉市人力资源管理部分写一篇建议报告,要求论点明确,论据充分。

二、模型假设

- (1) 就写假设就行了吧
- (2)
- (3)
- (4)

三、符号说明

符号	说明	单位
C_0	污染源初始浓度	单位
$C(x, t)$	污染浓度随时空变化	单位
u_x	江河平均纵向流速	单位
E_x	铊在江河纵向弥散系数	单位

四、问题一模型的建立与求解

4.1 问题的描述与分析

针对问题一，本题要求对武汉的人才吸引力做出量化评价。对于人才吸引力水平模型，现在主要存在的问题是定性研究较多、量化评价较少，对变量之间的关系讨论更是不足。本文采用因子分析法建立模型，将通过不同的因子对变量的影响程度进行分析，建立客观的评价模型。首先根据题目要求和以往人才吸引要素研究，寻找合适的人才吸引力因素，确立要素框架并寻找各个因素相对应的数据。在将所得数据进行无量纲化处理后，进行 KMO 检验和巴特利特检验以验证所选变量是否适合做因子分析。通过主成分分析法确定公因子数目，求解旋转后因子荷载矩阵并对公因子命名，最后在对武汉各年人才吸引力做出综合得分，其流程图如图所示：



图 1 因子分析算法流程图

4.2 预备工作

4.2.1 评价指标体系构建

参考国家统计局对人才标准的划分，此处对人才的界定是指，具有大专及以上学历的人员，以及具有初级以上专业技术职称的人员或在专业技术岗位上工作的人员。

在构建武汉市人才吸引力评价指标体系时，在指标方面，若选取总量太多，在评价模型时太过复杂，可行性低，若数据太少则评价不够准确，脱离实际性。因此为尽可能客观反映吸引力，我们在遵循建立指标体系的科学性、系统性、可操作性、可比性原则[0]，并兼顾总量指标和相对指标，选取城市发展前景、主要行业收入、政府影响、环境因素、年末总人口共 5 个二级指标和 33 个三级指标，多维度、全方面的考虑各个指标对人才吸引力的影响。具体指标体系如图 2 所示：



图2 城市人才吸引力指标体系

4.2.2 同向化处理

在评价城市人才的指标中，工业废水废气排放量和商品房平均销售价格不是越高越好，为方便比较，对这三个指标进行同向化处理 [0]，其公式如下：

$$\tilde{X} = -(X - \bar{X}) \quad (1)$$

式中：X 原始指标， \bar{X} 为原始指标 X 的平均值， \tilde{X} 为同向化指标。

4.2.3 标准化处理

为了对变量进行比较并消除由于观测量纲的差异及数量级差异所造成的影响, 将样本观测数据进行标准化处理，使标准化的变量均值为 0，方差为 1。其处理方法如下：

$$Z = \frac{X - \bar{X}}{\delta_X} \quad (2)$$

式中： \bar{X} 为 X 的均值， δ_X 是 X 的标准差。

4.2.4 KMO 与 Bartlett 检验

根据标准化处理后的指标数据，利用 SPSS23.0 统计软件进行 KMO 和巴特利特检验，以确认所选变量是否适合做因子分析，结果如表所示：

表 1 KMO 和 Bartlett 的检验

取样足够度的 Kaiser-Meyer-Olkin 度量		0.859
Bartlett 的球形度检验	近似卡方	1183.981
	df	78
	Sig.	.000

由上表可知，巴特利特球度检验统计量的观测值为 1183.981，相应的概率 P 值接近 0。若显著性水平 为 0.5，概率 P 小于显著性水平，应拒绝零假设，认为相关系数矩阵与单位阵有显著差异，即因子协方差矩阵不是单位阵。同时， KMO 值为 0.859， $KMO > 0.8$ ，根据 KMO 度量标准可知原有变量适合进行因子分析。

4.3 模型的建立与求解

4.3.1 公因子的确定

根据上述指标标准化后，其对应相关系数矩阵部分数据见表 2 所示（详见附录）：

表 2 相关系数矩阵部分数据

	X_1	X_2	X_3	X_4	X_{33}
X_1	1.000	0.727	-0.423	0.967	-0.928
X_2		1.000	-0.796	0.822	-0.814
X_3			1.000	-0.620	0.682
X_4				1.000	-0.955
.....				
X_{33}					1.000

设 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$ 为样本相关系数矩阵 R 的特征值， $\eta_1, \eta_2, \cdots, \eta_p$ 为相应的标准正交化特征向量。设 $m < p$ ，则因子载荷矩阵 Λ 为：

$$\Lambda = \left[\sqrt{\lambda_1} \eta_1, \sqrt{\lambda_2} \eta_2, \cdots, \sqrt{\lambda_m} \eta_m \right] \quad (3)$$

用 $\mathbf{R} - \mathbf{\Lambda}\mathbf{\Lambda}^T$ 对角元来估计特殊因子的方差:

$$\sigma_i^2 = 1 - \sum_{j=1}^m \alpha_{ij}^2 \quad (4)$$

式中: σ_i^2 为方差, α_{ij}^2 表示载荷因子。由上述分析, 得总方差解释部分数据如表 3 所示:

表 3 总方差解释

初始特征值				提取载荷平方和		
成分	总计	方差百分比	累积	总计	方差百分比	累积
0	25.330	76.758	76.758	25.330	76.758	76.758
1	4.025	12.197	88.956	4.025	12.197	88.956
2	2.373	7.191	96.147	2.373	7.191	96.147
3	1.271	3.852	99.920	1.271	3.852	99.920
4	5.129e-15	1.554e-14	100.0			
5	1.330e-15	4.031e-15	100.0			
6	6.062e-16	1.837e-15	100.0			
7	5.067e-16	1.535e-15	100.0			
8	4.532e-16	1.373e-15	100.0			
9	3.335e-16	1.011e-15	100.0			
10	3.137e-16	9.506e-16	100.0			
11	2.585e-16	7.834e-16	100.0			
12	1.990e-16	6.031e-16	100.0			
.....	100.0			
.....	100.0			
33	-2.279e-15	-6.908e-15	100.0			

由表 3 特征根知, 因子 F_1 的特征值 $\lambda_1 = 25.330$, 占方差的 76.75%。由碎石图 3 可知, 当提取 1, 2 个公因子时, 特征值变化非常明显, 当提取第 5 个以后的公因子时, 特

征值变化比较小，基本趋于平缓。由此说明，提取 4 个公因子对原变量信息的刻画有显著作用。因此，在这里我们提取 4 个公共因子，这 4 个公因子的累计方差达到 99.92%，即这 4 个公因子可以反映原来 33 个指标的 99.92% 的信息量，可见采用前 4 个公因子对武汉 2013—2017 年人才吸引力进行评价是比较合适的。



图 3 碎石图

4.3.2 因子命名与载荷矩阵计算

表 4 是初始因子载荷矩阵（详见附录），由此可写出因子分析模型的如下：

$$\begin{aligned}
 X_1 &= 0.989F_1 + 0.123F_2 - 0.074F_3 - 0.031F_4 \\
 X_2 &= 0.795F_1 - 0.556F_2 - 0.193F_3 - 0.151F_4 \\
 &\dots\dots\dots \\
 X_{33} &= -0.958F_1 + 0.176F_2 - 0.060F_3 + 0.220F_4
 \end{aligned}
 \tag{5}$$

表 4 初始因子载荷矩阵

指标	F_1	F_2	F_3	F_4
地方财政预算支出	0.989	0.123	-0.074	-0.031
固定资产总额	0.795	-0.556	-0.193	0.151
年末邮政局数	-0.548	0.768	-0.333	0.008
执业医师人数	0.994	-0.027	-0.245	0.002
.....
年末固定电话数	-0.958	-0.176	-0.060	0.220

表 4 中的每个数据表示了相应因子变量对相应原变量的相对重要程度。由于得到的公共因子与各指标的载荷分布归类比较困难，需要对因子载荷矩阵进行正交旋转，这里运用方差最大正交旋转法 [0]，得到旋转后的因子载荷矩阵表 5 所示（详见附录）：

表 5 旋转后的因子载荷矩阵

指标	F_1	F_2	F_3	F_4
地方财政预算支出	0.870	0.356	0.076	0.331
固定资产总额	0.469	0.311	0.678	0.473
年末邮政局数	-0.291	-0.390	-0.867	0.105
执业医师人数	0.787	0.496	0.259	0.259
.....
年末固定电话数	-0.872	-0.277	-0.384	-0.124

根据表 5 发现，旋转在 9 次迭代后已收敛。旋转后的因子系数已经明显向两极分化，有了更鲜明的实际意义。因子载荷的绝对值越大，则表明该因子与变量的重叠性越高，在解释因子的时候就越重要。

第一因子 F_1 主要包含地区生产总值、第一第二产业生产总值、在岗职工平均工资和人均 GDP 等，这些指标包含了地区工业发展水平以及人均薪酬水平故将其命名为工业发展与薪酬因子。第二因子 F_2 主要包含持证医师人数，医院卫生院个数以及一年内

空气质量达到及好于二级的天数等，这些指标包含了地区医疗卫生水平和居民生活环境，故将其命名为**医疗卫生环境因子**。第三因子 F_3 主要包含固定资产投资总额与货物进出口总额，该指标反映了城市贸易与经济发展水平，故将其命名为**经济贸易因子**。第四因子 F_4 包含年末总人口与旅游生产总值，反映了城市人口与拥挤程度故将其命名人口与拥挤程度因子。

4.3.3 求得因子得分和综合绩效得分

采用回归法估计因子得分系数，并输出因子得分系数矩阵，其结果见表 6 所示（详见附录）：

表 6 成分得分系数矩阵

指标	F_1	F_2	F_3	F_4
地方财政预算支出	0.052	-0.017	-0.021	0.024
固定资产总额	-0.063	0.045	0.127	0.137
年末邮政局数	0.022	-0.079	-0.195	0.109
执业医师人数	0.007	0.064	0.021	0.009
.....
年末固定电话数	-0.109	0.093	-0.068	0.094

利用 SPSS 软件计算得到旋转后的因子贡献及贡献率如表 7 所示：

表 7 旋转载荷平方和

因子	贡献	贡献率	累计
F_1	18.238	55.267	55.267
F_2	5.889	17.845	73.112
F_3	4.826	14.623	87.735
F_4	40.470	12.265	99.920

由此计算出的因子得分，可以量化描述城市人才吸引力水平，利用因子得分可以从不同角度对城市人才吸引力水平进行比较分析。为了便于对各城市进行人才吸引力评

价，现利用武汉每年的因子得分表计算综合得分，吸引力水平的获取是基于总方差分解表中旋转后各因子的方差贡献率及计算所得的城市各因子得分获取的。其计算公式如下：

$$F = (55.267F_1 + 17.845F_2 + 14.623F_3 + 12.265F_4) / 99.92 \quad (6)$$

经计算，纵向比较武汉每年的综合因子得分后绘制直方图：

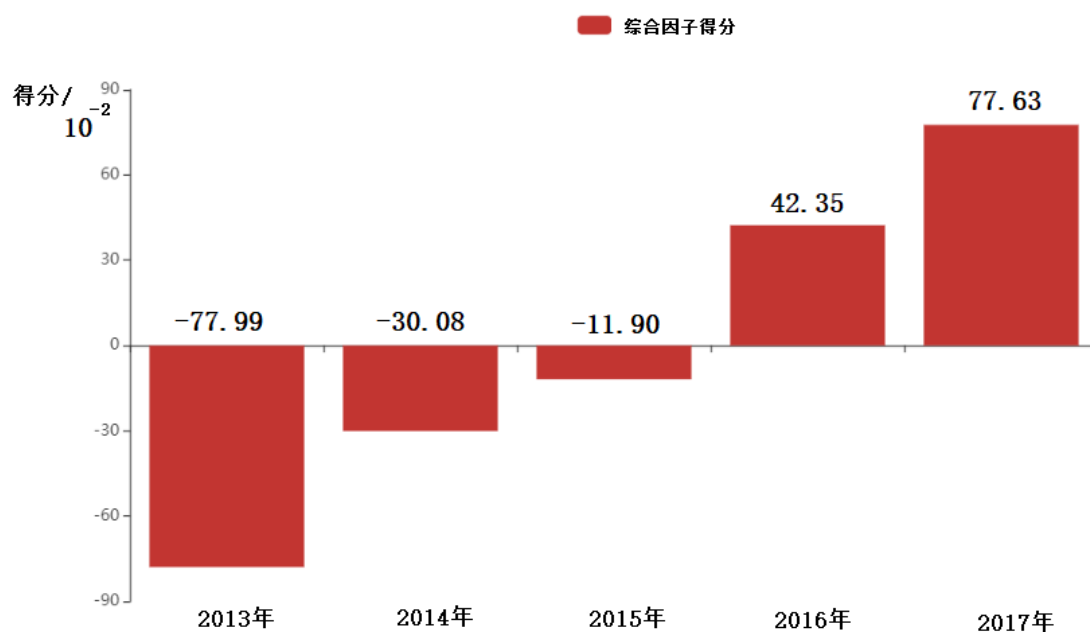


图 4 综合因子得分

4.4 结果分析

五、 问题二模型的建立与求解

5.1 问题的描述与分析

问题二要求针对具体人才类别，量化分析比较武汉市与其它同类城市的人才吸引力。本文在问题一模型的基础上改进，将人才分为三类——第一产业人才、第二产业人才和第三产业人才，并将问题一中的人才吸引力因素通过灰色相关度分成针对各类型人才的主要影响因素，再对分类后数据进行因子分析得到成都、天津、西安、南京和武汉 2013 2017 年分别对于三类人才吸引力的因子分析模型，并计算每个城市每年针对于三类人才的人才吸引力综合得分。

5.2 模型的建立与求解

5.2.1 人才及其吸引力影响因素的分类

不同产业的人才会有不同的关注焦点，所以城市对不同产业类型的人才吸引力影响因素也存在差异。我们把人才分为三个类型——第一产业人才、第二产业人才和第三产业人才。通过查阅资料收集了武汉 2013 2017 年第一、第二和第三产业大专以上学历从业人员人数，在将所收集数据与问题一中的人才吸引力要素标准化得到表 (8.1)

(表 8.1) 利用灰色关联度理论分别分析三类人才的 33 个影响因素，得到每个影响因子与用电量之间的关联度系数，具体步骤如下：确定因素序列和特征序列。本文分别以武汉 2013 2017 年第一、第二和第三产业大专以上学历从业人员人数作为特征序列，设为 $x_k(p)$ ，采用 $n(n = 15)$ 个数据： $x_k(p) = x_k(1), x_k(2), \dots, x_k(5)(k = 1, 2, 3)$ ；将 33 个人才吸引要素作为因素序列，设为 $x_i(t)$ ，其中有 $m(m = 33)$ 个子序列，每个子序列对应 5 个数据： $x_i(p) = x_i(1), x_i(2), \dots, x_i(5)$ 。

计算两极最小差 k 和最大差 K 。计算出特征序列与因素序列的序列差为 $\Delta_i(t)$ ，找到结果中的最小差和最大差。其中：

$$\begin{aligned}\Delta_1(t) &= |x_0(t) - x_1(t)| \\ k &= \min \min \Delta_i(t) \\ K &= \max \max \Delta_i(t)\end{aligned}\tag{7}$$

灰色关联系数和关联度。因素序列和特征序列在第 T 点的关联系数为：

$$\Phi_{0,(p)} = \frac{k + \varepsilon K}{\Delta_{0i}(t) + \varepsilon K}\tag{8}$$

利用 MATLAB 软件分析计算得出湖北省年用电量与 33 个影响因素的相关联程度如表 (8.3) 所示。

5.2.2 分类型人才吸引力分析

将归类后所得的影响第一产业人才吸引力的 33 个影响因素在五个城市中 2013 2017 年的数据，使用 SPSS 进行因子分析，按照累积贡献率的原则提取公因子，以综合性指标来相对全面的反映出全体因子对第一产业人才吸引力的影响情况。按照主因子的提取原则，很明显地看出可以用 4 个主因子来描述此 33 个因子的影响。然后得出主因子的解释方差解释如表 9 所示：

表 (9) 通过表 9 中的方差贡献率及计算所得的城市各因子得分得到各城市每年的第一产业人才吸引力综合得分。其计算公式如下：

$$F = (55.267F_1 + 17.845F_2 + 14.623F_3 + 12.265F_4) / 99.92\tag{9}$$

将得到的五个城市 2013 2017 年的综合得分绘制为折线图的到图 (1) 如下所示图 (1) 同理将归类后所得的影响第二产业人才吸引力的 □□ 个影响因素使用 SPSS 进行因子分析, 得到方差解释表如表 10 所示 (表 10) 得到第二产业人才吸引力综合得分计算公式:

$$F = (55.267F_1 + 17.845F_2 + 14.623F_3 + 12.265F_4) / 99.92 \quad (10)$$

第二产业人才吸引力综合得分折线图: 图 (2)

同理第三产业人才吸引力的 □□ 个影响因素因子分析得得方差解释表如表 11 所示 (表 11)

高级人才吸引力综合得分计算公式:

$$F = (55.267F_1 + 17.845F_2 + 14.623F_3 + 12.265F_4) / 99.92 \quad (11)$$

高级人才吸引力综合得分折线图:

5.3 结果分析

由图表可以看出对 XX 人才吸引力逐年上涨... 由人才吸引力这一简单的方面可以反映我国西部和东部的综合实力差异和自然环境造成的优劣。可以说, 人才吸引力的评价就是综合实力评价的缩影版。

具体浓度分布如图 ?? 所示:

六、 问题三的求解

6.1 写给武汉市人力资源部门的建议报告

武汉市人力资源管理部分的领导:

武汉市作为全国科技和教育事业最发达的地区之一, 无论在拥有高校数量还是毕业生质量方面都位居前列。就业政策体系不断完善, 扶持就业创业力度不断加大, 社会保障体系建设取得突破性进展, 人才队伍建设卓有成效。随着” 黄鹤英才计划”, ” 3551 人才计划” 等重大人才工程的推进, 累计引进, 培养中央” 千人计划”、国家” 万人计划”、省” 百人计划” 等顶尖人才 375 人, 吸纳引进海内外高层次创新创业人才 4000 余人, 数量居中部城市第一。

人才决定着一个城市的发展前景, 但城市必须拥有足够的吸引力才能让人才流入。人才在选择移居城市时必然会去对这个城市进行量化评估并与其他城市进行对比, 进而去选择最适和发展的城市。发展前景, 迁移成本、薪资待遇以及城市各方面环境因素、政府策略都是会是考虑的因素。我们只有紧紧抓住人才这个关键环节, 才能在新一轮经济发展和城市竞争中取得优势, 因此我们建立了人才吸引力影响因素和政策效力的模型, 在政策制定有很大的理论和现实意义, 并向贵部门提出相关建议。

通过以 2017 年国家统计局年鉴和 2017 年各大城市统计年鉴和各式人力资源和社会保障官网的相关数据为依据,对以武汉市在内的同类型城市(西安、成都、天津、南京、武汉)人才吸引力水平进行分析。我们结合了因子分析法和主成分分析法地代数模型进行研究,首先构建评价指标体系。遵循客观地、科学性和系统性地原则选取了城市发展前景、主要行业收入、政府影响、环境因素、年末总人口共 5 个二级指标和 33 个三级指标。并应用因子分析将 33 个三级指标综合成 4 个公共因子,这 4 个公因子的累计方差达到 99.2%。用这 4 个公因子来对武汉城市吸引力进行评价,再利用主成分分析法提取,构建未转轴的因子载荷矩阵,得到因子的分析模型后,进行正交旋转,对两级分化的因子系数进行命名,第一个主因子包含 xxxxxxxx,第二个主因子包含 xxxxxxxxxx,第三个主因子包含 xxxxxxxxxx,第四个主因子包含 xxxxxxxxxx。多次旋转迭代后的成分矩阵已收敛,得到最终的人才吸引力评价模型,计算各因子的作用权重,并给出五座城市的综合排名。研究数据表明.....

在本次研究中,武汉市的城市基础设施建设、经济发展状况.....

因此,针对以上情况我们向有关领导提出以下建议

1. 促进城市经济发展,增加城市提供就业机会能力,大力实施大众创业万众创新战略,着力打造创业梦想之城。人才被城市的经济发展水平及其提供的发展前景所吸引。发展经济是吸引人才的基础条件,要从区域、产业、企业三个层次促进经济的发展。依靠核心城市群的带动作用,加强区域协调发展,统筹城乡就业工作,推动生产要素在区域内自由流和优化配置。解决结构性就业矛盾,把产业规划和人才引进规划匹配起来。企业作为吸纳人才的主体,要为企业的发展提供宽松稳定的环境与高效清明的服务环境,对企业实行税收上的减免与政策上的扶持,增强企业的实力及其对人才的吸引能力。

2. 改善公共服务与自然环境,加强城市舒适性建设。从满足不同层次需求的角度来说,高层次的需求是在经济水平之上的更加舒适的生活环境。加强城市规划、加快基本公共服务供给侧结构改革。建立更加公平可持续的社会保障制度,着力构建社会和谐之基。进一步推进社会保障制度改革,建立健全社会保障待遇确定和正常调整机制。建立健全社会保险市级统筹制度,完善筹资机制。继续做好社会保险扩面工作,要以人为本,促进社会和谐发展,增强城市对于生活质量要求较高人才的吸引力。

3. 优化城市人才政策,增强城市为人才提供就业服务的能力。着力于打造国际性创行人才高地。不断优化人才创新创业环境,创新工作体制机制。积极打造国际人才自由港,努力组建规模宏大、具有较强国际竞争力、引领支撑国家创新中心建设的人才队伍。发挥市场调节供需关系的作用。充分尊重人才的成长规律,人才的评价与激励涵盖人才的品德等方面,同时尊重市场的规律,实现人才资源的最优组合。更注重对人才的人文关怀和人才个人价值体现,从而体现以人为本的发展理念。

本研究还存在一些局限性,本研究的数据均来源于各类统计年鉴和人力资源和社会保障厅官网,受数据的限制,尚有其他相关重要的测量指标数据未能获取,仅能得到相

对条件下的相对结论，使得本研究与现实情况之间出现一些偏差，未来或需引入一些新的能够容纳更多信息的方法。

七、模型的评价

7.1 模型的优点

XXXXXXXXXXXXXXXXXXXXXXXXXXXX

7.2 模型的缺点

XXXXXXXXXXXXXXXXXXXXXXXXXXXX

7.3 模型的改进与展望

XXXXXXXXXXXXXXXXXXXXXXXXXXXX

附录 A 代码

A.1 数据预处理–python 源代码

```
from sklearn.preprocessing import StandardScaler
from sklearn.decomposition import PCA
from factor_analyzer import FactorAnalyzer

#导入数据
df = pd.read_excel("wut.xls")
# print(df.head(5))

#转置
df = pd.DataFrame(df.values.T, index=df.columns, columns=df.index)
# print(df['商品房平均销售价格（元/平方米）'])

数据正向处理
df['商品房平均销售价格（元/平方米）'] =
    -(df['商品房平均销售价格（元/平方米）']-sum(df['商品房平均销售价格（元/平方米）'])/5)
df['工业废水排放量（万吨）'] = -(df['工业废水排放量（万吨）']-sum(df['工业废水排放量（万吨）'])/5)
df['工业二氧化硫排放量（吨）'] =
    -(df['工业二氧化硫排放量（吨）']-sum(df['工业二氧化硫排放量（吨）'])/5)
print(df['商品房平均销售价格（元/平方米）'])
print(df.head(5))
df.to_excel("武汉.xls")

#数据标准化
df = StandardScaler().fit_transform(df)
# print(df)
df=pd.DataFrame(df)
df.to_excel("武汉.xls")
#
#主成分分析
pca = PCA(n_components=4)
newX = pca.fit_transform(df)
print(newX)

#返回所保留的n个成分各自的方差百分比
print("它代表降维后的各主成分的方差值占总方差值的比例，这个比例越大，则越是重要的主成分")
print(pca.explained_variance_ratio_)
print("它代表降维后的各主成分的方差值，方差值越大，则说明越是重要的主成分")
print(pca.explained_variance_)
```

A.2 数据可视化–python 源代码

```

from pyecharts import options as opts
from pyecharts.charts import Bar
from pyecharts.globals import ThemeType

def bar_xyaxis_name() -> Bar:
    c = (
        Bar(init_opts=opts.InitOpts(theme=ThemeType.WHITE))
        .add_xaxis(['2013年', '2014年', '2015年', '2016年', '2017年'])
        .add_yaxis("得分", [-77.99, -30.08, -11.90, 42.35, 77.63])
        .set_global_opts(
            title_opts=opts.TitleOpts(title="综合因子得分"),
            # yaxis_opts=opts.AxisOpts(name="Score"),
            toolbox_opts=opts.ToolboxOpts(),
            # xaxis_opts=opts.AxisOpts(name="Year"),

            brush_opts=opts.BrushOpts(),
        )
        # .add(is_datazoom_show=True)
    )
    return c

b = bar_xyaxis_name()
b.render()

```

表 8 相关系数矩阵

1.000	0.727	-0.423	0.967	0.973	0.937	0.799	0.989	0.989	-0.988	0.994	0.969
0.727	1.000	-0.796	0.822	0.832	0.721	0.760	0.707	0.802	-0.657	0.716	0.746
-0.423	-0.796	1.000	-0.620	-0.470	-0.293	-0.825	-0.469	-0.518	0.311	-0.357	-0.586
0.967	0.822	-0.620	1.000	0.947	0.853	0.909	0.961	0.973	-0.919	0.941	0.985
0.973	0.832	-0.470	0.947	1.000	0.969	0.752	0.951	0.986	-0.961	0.981	0.925
0.937	0.721	-0.293	0.853	0.969	1.000	0.608	0.919	0.947	-0.962	0.965	0.851
0.799	0.760	-0.825	0.909	0.752	0.608	1.000	0.838	0.828	-0.718	0.735	0.917
0.989	0.707	-0.469	0.961	0.951	0.919	0.838	1.000	0.987	-0.979	0.976	0.985
0.989	0.802	-0.518	0.973	0.986	0.947	0.828	0.987	1.000	-0.973	0.982	0.972
-0.988	-0.657	0.311	-0.919	-0.961	-0.962	-0.718	-0.979	-0.973	1.000	-0.995	-0.935
0.994	0.716	-0.357	0.941	0.981	0.965	0.735	0.976	0.982	-0.995	1.000	0.940
0.969	0.746	-0.586	0.985	0.925	0.851	0.917	0.985	0.972	-0.935	0.940	1.000
0.960	0.785	-0.629	0.996	0.922	0.827	0.933	0.965	0.963	-0.911	0.926	0.994
0.797	0.318	0.034	0.720	0.704	0.671	0.473	0.731	0.700	-0.794	0.798	0.705
0.898	0.465	-0.179	0.854	0.806	0.750	0.665	0.858	0.826	-0.881	0.885	0.851
0.246	-0.478	0.493	0.094	0.048	0.142	0.025	0.279	0.134	-0.321	0.241	0.217
-0.670	-0.615	0.016	-0.578	-0.771	-0.801	-0.208	-0.570	-0.659	0.691	-0.732	-0.491
-0.898	-0.674	0.164	-0.797	-0.940	-0.978	-0.489	-0.850	-0.891	0.928	-0.939	-0.772
0.614	0.436	-0.428	0.713	0.512	0.326	0.712	0.573	0.551	-0.515	0.553	0.668
-0.119	-0.603	0.873	-0.285	-0.211	-0.114	-0.537	-0.201	-0.254	0.048	-0.073	-0.283
0.914	0.422	-0.213	0.831	0.816	0.832	0.702	0.939	0.876	-0.938	0.904	0.899
-0.927	-0.690	0.597	-0.970	-0.854	-0.745	-0.939	-0.937	-0.913	0.872	-0.882	-0.979
0.985	0.813	-0.569	0.995	0.971	0.899	0.876	0.978	0.990	-0.950	0.967	0.985
0.991	0.793	-0.529	0.982	0.979	0.930	0.849	0.990	0.999	-0.970	0.980	0.982
0.513	0.930	-0.674	0.585	0.685	0.624	0.493	0.486	0.615	-0.467	0.529	0.495
0.766	0.762	-0.758	0.895	0.726	0.540	0.933	0.757	0.769	-0.662	0.705	0.854
0.967	0.699	-0.534	0.981	0.905	0.822	0.903	0.972	0.952	-0.929	0.935	0.992

A.3 相关系数矩阵

A.4 初始因子载荷矩阵

表5：成分矩阵 ^a				
	成分			
	1	2	3	4
Zscore(地方财政预算内支出（万元）)	.989	.123	-.074	-.031
Zscore(固定资产投资总额（万元）)	.795	-.556	-.193	.151
Zscore(年末邮政局（所）数（处）)	-.548	.768	-.333	.008
Zscore(执业（助理）医师（人）)	.994	-.068	.071	.056
Zscore(货运量（万吨）)	.969	-.027	-.245	.002
Zscore(年末总人口（万人）)	.902	.090	-.397	-.144
Zscore(客运量（万人）)	.870	-.279	.402	-.058
Zscore(第一产业（亿元）)	.983	.091	-.003	-.159
Zscore(地区生产总值（亿元）)	.991	-.013	-.106	-.086
Zscore: 商品房平均销售价格（元/平方米）	-.958	-.211	.158	.116
Zscore: 人均GDP（元/人）	.972	.153	-.174	-.035
Zscore(工业废水排放量（万吨）)	.987	-.002	.131	-.092
Zscore(第三产业占比)	.988	-.051	.148	.004
Zscore(房地产和投资业)	.745	.578	.013	.334
Zscore(工业和建筑业)	.869	.430	.103	.221
Zscore(交通业)	.152	.885	.303	-.318
Zscore(旅游业)	-.621	-.116	.700	-.334
Zscore(邮政业)	-.849	-.177	.497	-.013
Zscore(农业)	.659	.127	.510	.538
Zscore(地方财政预算内收入（万元）)	-.233	.899	-.159	.335

Zscore(地方财政预算内收入(万元))	-.233	.899	-.159	.335
Zscore(剧场、影剧院(个))	.873	.383	.081	-.290
Zscore(经济增长率)	-.956	-.031	-.292	-.026
Zscore(社会消费品零售总额(万元))	.999	-.031	-.005	.009
Zscore(第三产业(亿元))	.996	-.002	-.056	-.068
Zscore(货物进出口总额(万美元))	.571	-.672	-.463	.094
Zscore(医院、卫生院(个))	.841	-.222	.390	.302
Zscore(年末实有公共(汽)电车营运车辆数(辆))	.981	.081	.178	-.007
Zscore(普通高等学校在校学生数(人))	-.984	-.003	-.168	.064
Zscore(在岗职工平均工资(元))	.993	-.048	-.061	-.089
Zscore(工业二氧化硫排放量(吨))	.946	.196	.163	-.199
Zscore(第二产业(亿元))	.970	-.041	-.211	-.118
Zscore(空气质量达到及好于二级的天数(天))	.931	-.202	.078	.294
Zscore(年末固定电话用户数(万户))	-.958	.176	-.060	.220

A.5 旋转后的因子载荷矩阵

旋转后的成分矩阵 ^a				
	成分			
	1	2	3	4
Zscore(地方财政预算内支出(万元))	.870	.356	.076	.331
Zscore(固定资产投资总额(万元))	.469	.311	.678	.473
Zscore(年末邮政局(所)数(处))	-.291	-.390	-.867	.105
Zscore(执业(助理)医师(人))	.787	.496	.259	.259
Zscore(货运量(万吨))	.807	.268	.210	.484
Zscore(年末总人口(万人))	.850	.054	.091	.515
Zscore(客运量(万人))	.692	.546	.462	-.100
Zscore(第一产业(亿元))	.922	.303	.121	.208
Zscore(地区生产总值(亿元))	.870	.295	.213	.333
Zscore: 商品房平均销售价格(元/平方米)	-.907	-.237	.010	-.348
Zscore: 人均GDP(元/人)	.866	.289	.041	.406
Zscore(工业废水排放量(万吨))	.870	.425	.209	.135
Zscore(第三产业占比)	.811	.501	.249	.169
Zscore(房地产和投资业)	.582	.574	-.449	.361
Zscore(工业和建筑业)	.709	.594	-.266	.271

Zscore(交通业)	.477	.043	-.793	-.376
Zscore(旅游业)	-.384	-.099	.044	-.917
Zscore(邮政业)	-.748	-.087	.022	-.658
Zscore(农业)	.304	.951	-.031	.038
Zscore(地方财政预算内收入(万元))	-.165	.076	-.960	.214
Zscore(剧场、影剧院(个))	.959	.222	-.171	.039
Zscore(经济增长率)	-.790	-.588	-.165	-.052
Zscore(社会消费品零售总额(万元))	.825	.423	.226	.300
Zscore(第三产业(亿元))	.867	.339	.203	.303
Zscore(货物进出口总额(万美元))	.288	.019	.744	.602
Zscore(医院、卫生院(个))	.499	.782	.365	.078
Zscore(年末实有公共(汽)电车营运车辆数(辆))	.840	.512	.120	.135
Zscore(普通高等学校在校学生数(人))	-.854	-.465	-.202	-.117
Zscore(在岗职工平均工资(元))	.866	.318	.249	.295
Zscore(工业二氧化硫排放量(吨))	.933	.358	.020	.040
Zscore(第二产业(亿元))	.863	.203	.236	.398
Zscore(空气质量达到及好于二级的天数(天))	.586	.636	.355	.355
Zscore(年末固定电话用户数(万户))	-.872	-.277	-.384	-.124

A.6 成分得分系数矩阵

	成分			
	1	2	3	4
Zscore(地方财政预算内支出(万元))	.052	-.017	-.021	.024
Zscore(固定资产投资总额(万元))	-.063	.045	.127	.137
Zscore(年末邮政局(所)数(处))	.022	-.079	-.195	.109
Zscore(执业(助理)医师(人))	.007	.064	.021	.009
Zscore(货运量(万吨))	.031	-.042	.011	.096
Zscore(年末总人口(万人))	.093	-.160	-.009	.091
Zscore(客运量(万人))	.035	.077	.084	-.150
Zscore(第一产业(亿元))	.100	-.072	-.002	-.050
Zscore(地区生产总值(亿元))	.066	-.057	.016	.015
Zscore: 商品房平均销售价格(元/平方米)	-.090	.085	.037	-.020
Zscore: 人均GDP(元/人)	.055	-.044	-.029	.057
Zscore(工业废水排放量(万吨))	.068	-.003	.017	-.069
Zscore(第三产业占比)	.028	.054	.022	-.038
Zscore(房地产和投资业)	-.076	.205	-.160	.131

Zscore(工业和建筑业)	-.036	.165	-.113	.059
Zscore(交通业)	.178	-.093	-.185	-.230
Zscore(旅游业)	.102	-.027	.058	-.378
Zscore(邮政业)	-.035	.097	.044	-.186
Zscore(农业)	-.186	.432	-.060	.040
Zscore(地方财政预算内收入（万元）)	-.091	.151	-.247	.179
Zscore(剧场、影剧院（个）)	.164	-.123	-.063	-.132
Zscore(经济增长率)	-.022	-.101	-.002	.080
Zscore(社会消费品零售总额（万元）)	.028	.020	.014	.017
Zscore(第三产业（亿元）)	.060	-.034	.013	.005
Zscore(货物进出口总额（万美元）)	-.053	-.056	.154	.206
Zscore(医院、卫生院（个）)	-.105	.273	.042	-.007
Zscore(年末实有公共（汽）电车营运车辆数（辆）)	.039	.056	-.009	-.053
Zscore(普通高等学校在校学生数（人）)	-.057	-.022	-.014	.072
Zscore(在岗职工平均工资（元）)	.066	-.048	.025	-.002
Zscore(工业二氧化硫排放量（吨）)	.120	-.053	-.023	-.123
Zscore(第二产业（亿元）)	.077	-.100	.024	.038
Zscore(空气质量达到及好于二级的天数（天）)	-.097	.195	.035	.099
Zscore(年末固定电话用户数（万户）)	-.109	.093	-.068	.094