

# “未来之城”的人口引力场建构—— 多元视角下雄安新区人口吸引力预测及因素探究

## 摘 要

在京津冀协同发展战略背景下，雄安新区作为重要枢纽，需要承接北京非首都功能疏解，因此，探究人口吸引力的影响因素及预测至关重要。

本研究通过百度智能云大数据、国家统计局及雄安官网等多平台，搜集目前最新数据。据此构建经济、人才、建设、生活**四维评价体系**对雄安新区建设至今的人口吸引力的影响因素展开深入研究并进行精确预测。

首先，立足于四个维度对研究调查的具体因素进行了详尽的描述性统计分析。之后，进行因子分析：**灰色关联分析**显示**经济（ $\Gamma=0.814$ ）、人才（ $\Gamma=0.746$ ）、生活（ $\Gamma=0.737$ ）、建设（ $\Gamma=0.660$ ）**四类因子与人口吸引力指数的灰色关联度递减。基于分析结果，建立 AR-PLSR 预测体系：**ARIMA (3,0,2)**模型基于 2020.12-2022.12 期间的人口吸引力指数数据**预测未来 6 个月**人口吸引力指数，结果为 **0.80、0.76、0.74、0.79、0.74、0.78**，得出未来一段时间内指数应呈波动上升状态。**偏最小二乘回归（PLSR）模型**整合多维数据证实前 4 个潜在因子可解释近 80%的自变量信息（最终调整  **$R^2=0.802$** ），预测效度较好。

分析结果表明：雄安人口吸引力呈现“**经济主导（双轮驱动）、政策快响（雄才计划即时响应）、基建时滞（6-12 个月）**”特征。预测结果显示：新区人口吸引力将呈**波动上升趋势**。

具体地，经济因素为核心驱动；建设因素对人口吸引力的影响存在滞后效应，灰色关联度与偏最小二乘回归分析中的建设因素重要性（ $\Gamma=0.660$  &  $VIP<1$ ）均显著低于其余三类因素。特别地，人才因子中的“雄才计划”声量（ $VIP=1.229$ ）对人口吸引力起到较大推动作用。

**关键词：**雄安新区；人口吸引力；灰色关联分析；ARIMA 模型；PLSR 回归

# 目 录

一、引言.....	1
(一) 研究背景与意义.....	1
(二) 文献综述.....	2
1. 雄安新区人口吸引力相关理论.....	2
2. 雄安新区人口吸引相关预测模型.....	3
(三) 研究创新点.....	3
1. 多维度综合分析.....	3
2. 模型优化与持续学习.....	4
二、研究方法与数据概述.....	5
(一) 研究方法概述.....	5
1. 灰色关联分析.....	5
2. Arima 时序预测模型.....	6
3. 偏最小二乘回归模型.....	6
(二) 数据来源与变量说明.....	8
1. 人口吸引力.....	8
2. 多维度评价体系.....	8
(三) 缺失值处理.....	9
(四) 最大值归一化.....	10
三、雄安新区发展现状概述.....	11
(一) 经济因子.....	11
1. 经济活力.....	11
2. 就业机会.....	11
(二) 教育因子.....	12
1. 人才政策.....	12
2. 教育资源.....	13
(三) 建设因子.....	14
(四) 生活因子.....	14
四、人口吸引力多维度因素分类分析.....	16
五、雄安新区人口吸引力的预测分析.....	18
(一) Arima 时间序列预测模型.....	18
1. 模型参数选取.....	18

2. 残差检验.....	19
3. 预测结果.....	20
(二) 基于 PLSR (偏最小二乘回归) 回归预测模型 .....	20
1. 因子数量选取.....	21
2. 变量重要性探究.....	21
3. 预测结果.....	22
六、研究结论.....	23
七、提升雄安新区人口吸引力的策略建议.....	24
(一) 经济驱动：强化产业基础与就业质量.....	24
(二) 人才引育：优化政策效能与教育资源.....	24
(三) 建设提质：基础设施协同进步.....	24
(四) 生活提质：品质升级与生态价值转化.....	25
参考文献.....	26
附 录.....	28

表格与插图清单

表 1 多维评价体系.....9

表 2 部分数据展示.....10

表 3 灰色关联分析关联度.....16

表 4 AIC 数值变化.....18

表 5 潜在因子信息综合解释.....21

图 1 雄安新区地理位置..... 1

图 2 雄安新区周边情况.....2

图 3 多维度综合分析.....4

图 4 生活维度.....4

图 5 PLSR 模型介绍图 .....7

图 6 经济活力相关指标与吸引力散点矩阵图..... 11

图 7 2019-2024 年就业相关数据变化面积图 .....12

图 8 人才政策相关因素与吸引力随时间关系折线图..... 13

图 9 高校迁入声量趋势（分箱）与吸引力箱线图.....13

图 10 雄安新区各类 POI 数量变化对比棒棒糖图 .....14

图 11 生活品质相关因素与吸引力关系小提琴图..... 15

图 12 雄安新区人口吸引力-关联度比较分析 .....16

图 13 雄安新区人口吸引力-关联系数变化分析 .....17

图 14 雄安新区建设各阶段目标.....18

图 15 残差正态性检验 - QQ 图.....19

图 16 2020 - 12 月至 2022 - 12 月 ARIMA 时间序列预测 .....20

图 17 PLSR 模型的核心指标 .....21

图 18 2021 - 10 月至 2022 - 12 月 PLSR 预测 .....22

## 一、引言

### （一）研究背景与意义

在我国区域协调发展战略框架下，京津冀地区具有核心支点地位<sup>[1]</sup>。然而近年来，区域长期累积的结构性矛盾逐步显现。北京作为首都，承载了过多非首都功能，“大城市病”问题持续加剧，一系列发展难题对城市可持续发展形成显著制约。同时，京津冀区域内部发展不平衡、不协调的矛盾也尤为突出，河北部分地区与京津之间在经济发展水平、公共服务资源配置等方面存在较大差距。为着力破解发展困境，探寻人口经济密集区域优化开发新路径，优化京津冀城市布局与空间结构，培育创新驱动发展新动能，党中央、国务院于 2017 年 4 月 1 日作出设立雄安新区的重大战略部署。



图 1 雄安新区地理位置

在优化人口空间布局的框架下，雄安新区作为京津冀协同发展的重要枢纽，通过承接北京非首都功能疏解，正成为促进区域资源优化配置与协同发展的重要引擎。为了解决北京人口过于密集的状况，通过疏解非核心功能，有效促进人口向雄安新区的分阶段转移。

区域发展层面，雄安高起点规划、高标准建设，聚焦人工智能等高端高新产业，创造大量优质岗位，形成对国内外人才的强大吸引力。国家战略层面，

雄安通过营造宜居宜业环境，完善公共服务体系，探索新型人口治理模式，为各地在吸引、留住人才及提升人口素质方面提供经验，助力提升全国人口发展质量。



图 2 雄安新区周边情况

通过对雄安新区人口吸引力影响因素的探究，理论层面上，可深化对城市人口迁移规律的认知，推动全国人口发展模式优化<sup>[2]</sup>，实现人口与经济、社会、环境协调发展，提升整体人口发展质量。实践层面上，研究成果对雄安新区建设具有重要参考价值。通过系统梳理和分类关键影响因素，明确正向驱动与负向制约因素，可为政府部门制定精准化政策提供科学支撑。

与此同时，对于其他正在进行新区建设或致力于提升人口吸引力的地区而言，本研究成果也能提供宝贵的经验借鉴，助力其制定合理的人口发展策略，实现区域的可持续发展。

(二) 文献综述

1. 雄安新区人口吸引力相关理论

李国平,宋昌耀（2018）<sup>[3]</sup>提出为实现雄安新区的发展定位，雄安新区应当采取四大战略，从多方面入手，提升雄安新区对人口的吸引力，助力加快实施京津冀协同发展战略，为国内外新区新城建设提供宝贵的经验和示范。

王艳君,刘清滢,司丽丽（2022）<sup>[4]</sup>指出，影响雄安新区人口吸引力的因素包括政策导向、老龄化和性别比、教育资源、社会保障与基础设施等方面。并针对以上因素提出了针对性措施以提高雄安新区的生活质量和经济活力，从而增

强其对人口的吸引力。

孙明、吴敏(2023)<sup>[5]</sup>研究指出,雄安新区在交通、教育、医疗等基础设施建设方面的成果,提升了其对人口的吸引力。但与发达城市相比,雄安新区在优质公共服务资源的集聚和均衡分配上仍有提升空间。

## **2. 雄安新区人口吸引相关预测模型**

梁林,武晓洁(2020)<sup>[6]</sup>着重对雄安新区不同发展阶段的人才流动特征进行分析,识别出关键因素,构建适用于新区人才流动趋势预测的灰色预测模型,预测新区未来的人才流动趋势。

梁林,曾建丽,刘兵(2019)<sup>[7]</sup>参考深圳建设发展历程,同样通过构建不同时期的灰色系统预测模型分阶段预测了雄安新区的人口规模。

总体而言,现有较多研究对雄安新区整体规划、发展要素等理论进行探究,仅有少量文献对人口相关趋势进行预测分析,同时,研究方法相对单一,对新区各因素综合影响机制的研究仍显薄弱,难以精准剖析人口吸引力问题。

本研究深入挖掘新区各因素间的内在联系,从经济、人才、建设、生活四个维度全面深入地探究雄安新区人口吸引力问题,并创新融合机器学习研究方法,最终实现精确预测。

## **(三) 研究创新点**

### **1. 多维度综合分析**

现有关于雄安新区人口吸引力影响因素的研究,在指标构建方面存在明显短板,通常较为单一且局限,难以全面且深入地剖析复杂的人口吸引机制。本研究广泛收集已有数据,构建起一套全面且深入的多维度指标体系,从经济、人才、建设、生活四个维度,对雄安新区人口吸引力展开深度剖析。



图 3 多维度综合分析

首次将“央企迁入声量趋势”这类隐性因素纳入经济活力评估，拓展研究边界。结合“科技创新企业数量”和“企业数量”，采用能够体现发展潜力的经济指标，而非如“成熟”城市一般以单一 GDP 为导向。

在生活维度，考量全面内容丰富。聚焦群众关心的生活品质和环境因素，首次将“民众环保关注度”纳入分析，除此之外，囊括生活品质指数、休闲广场 POI 变化、公园 POI 变化以及对空气质量的检测分析，多方面深层建构。

### 生活维度与吸引力关系的多元因素

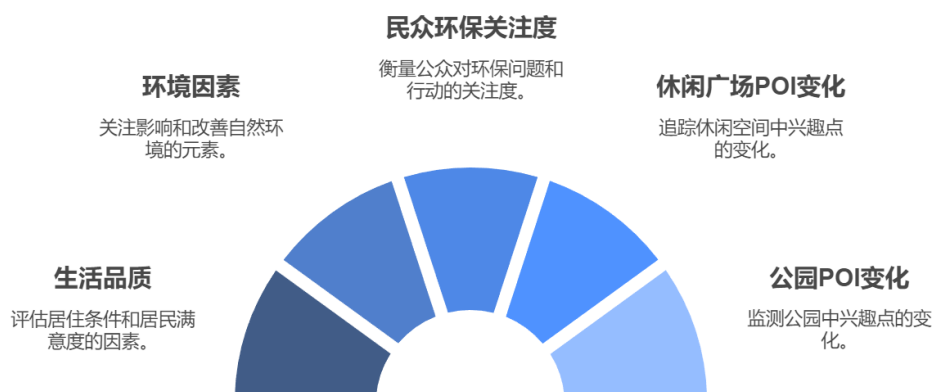


图 4 生活维度

## 2. 模型优化与持续学习

本研究构建了完善人口吸引力因素分析模型及 AR-PLSR 预测模型。

在描述性统计实现雄安新区人口吸引力基础特征图谱基础上，采用灰色关



联度模型量化因子水平；AR-PLSR 人口吸引力预测模型：首先使用 Arima 进行时序特征预测，之后使用 PLSR 的多维度因素特征预测，同时捕捉了时序及因素特征，并结合分析模型做出阐释，突破了传统人口预测中“单模型依赖”与“因素分析碎片化”的局限。

## 二、研究方法与数据概述

### （一）研究方法概述

本研究主要采用灰色关联分析、偏最小二乘法、Arima 时序预测三类模型，通过结构化数据与时序数据的融合分析，探究雄安新区人口吸引力的核心影响因素及未来趋势。

#### 1. 灰色关联分析

灰色关联分析<sup>[8]</sup>（Grey Relational Analysis, GRA）是一种基于灰色系统理论的多因素分析方法，用于评估各因素对系统行为的影响程度。它通过计算各因素与参考序列之间的关联度，确定各因素的重要性排序。该方法适用于数据量少、信息不完整的系统。

基本步骤：

##### ① 确定参考序列和比较序列

参考序列（目标序列）： $X_0 = (x_0(1), x_0(2), \dots, x_0(n))$

比较序列（因素序列）： $X_i = (x_i(1), x_i(2), \dots, x_i(n)), i = 1, 2, \dots, m$

##### ② 数据标准化

由于数据量纲不同，需进行标准化处理。

##### ③ 计算关联系数

关联系数反映比较序列与参考序列在各点的关联程度，计算公式为：

$$\gamma(x_0(k), x_i(k)) = \frac{\min_i \min_k |x_0(k) - x_i(k)| + \rho \max_i \max_k |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \rho \max_i \max_k |x_0(k) - x_i(k)|} \quad (1)$$

##### ④ 计算关联度

对各点的关联系数取平均，得到关联度：

$$\Gamma(X_0, X_i) = \frac{1}{n} \sum_{k=1}^n \gamma(x_0(k), x_i(k)) \quad (2)$$

##### ⑤ 排序分析

根据关联度大小，确定各因素对系统行为的影响程度。

## 2. Arima 时序预测模型

ARIMA (AutoRegressive Integrated Moving Average) 模型<sup>[9]</sup>是一种常用的时间序列预测方法，结合了自回归 (AR)、差分 (I) 和移动平均 (MA) 三个部分，适用于非平稳时间序列的预测。

### ⑥ AR (自回归)

利用时间序列过去值的线性组合预测未来值，公式为：

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t \quad (3)$$

其中， $p$  为自回归阶数， $\phi$  为系数， $\epsilon_t$  为误差。

### ⑦ I (差分)

通过对时间序列进行差分使其平稳，公式为：

$$\Delta y_t = y_t - y_{t-1} \quad (4)$$

### ⑧ MA (移动平均)

利用过去误差的线性组合预测未来值，公式为：

$$y_t = c + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} \quad (5)$$

其中， $q$  为移动平均阶数， $\theta$  为系数。

### ⑨ ARIMA 模型

将 AR、I、MA 结合，表示为 ARIMA( $p, d, q$ )，公式为：

$$\phi(L)(1-L)^d y_t = \theta(L)\epsilon_t \quad (6)$$

其中， $L$  为滞后算子， $\phi(L)$  和  $\theta(L)$  分别为 AR 和 MA 的滞后多项式。

## 3. 偏最小二乘回归模型

偏最小二乘回归<sup>[10]</sup> (Partial Least Squares Regression, PLSR) 是一种多元统计分析方法，主要用于处理多重共线性问题和高维数据。它通过提取自变量 (X) 和因变量 (Y) 之间的潜在变量 (Latent Variables) 来建立预测模型，特别适用于本研究自变量数量多且相关性强的场景。在降维的同时进行回归，避

免了多重共线性问题。

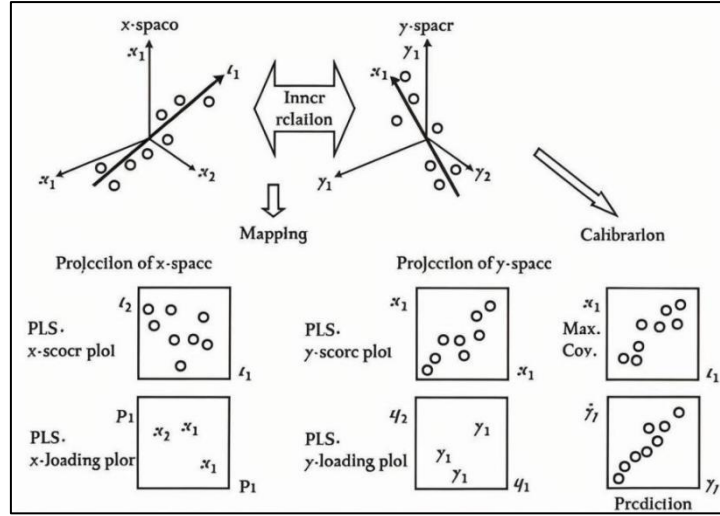


图 5 PLSR 模型介绍图

基本步骤：

### ① 数据准备

确定自变量矩阵  $X = (x_{ij})_{n \times p}$ （ $n$  为样本数量， $p$  为自变量个数）和因变量矩阵  $Y = (y_{ij})_{n \times q}$ （ $q$  为因变量个数）。

### ② 数据标准化

为消除量纲影响，对原始数据进行标准化处理，标准化后自变量矩阵记为  $E_0$ ，因变量矩阵记为  $F_0$ ，标准化公式为：

$$e_{ij} = \frac{x_{ij} - \bar{x}_j}{s_{x_j}}, f_{ij} = \frac{y_{ij} - \bar{y}_j}{s_{y_j}} \quad (7)$$

其中， $\bar{x}_j$ 、 $\bar{y}_j$  分别为  $X$ 、 $Y$  第  $j$  列均值， $s_{x_j}$ 、 $s_{y_j}$  分别为对应标准差。

### ③ 提取主成分

通过迭代提取主成分：

- (1) 计算  $E_0^T F_0 F_0^T E_0$  的最大特征值对应的特征向量  $u_1$ ，得到第一个主成分

$$t_1 = E_0 u_1 \quad (8)$$

- (2) 计算权重向量

$$p_1 = \frac{E_0^T t_1}{\|t_1\|^2}, q_1 = \frac{F_0^T t_1}{\|t_1\|^2} \quad (9)$$

- (3) 更新残差矩阵

$$E_1 = E_0 - t_1 p_1^T, F_1 = F_0 - t_1 q_1^T \quad (10)$$

重复上述步骤，得到后续主成分  $t_2, t_3, \dots, t_h$ ，直至满足停止条件（如累计方

差解释率达标)。

#### ④ 建立回归模型

提取  $m$  个主成分后，构建回归系数矩阵

$$B_m = (T_m^T T_m)^{-1} T_m^T F_0 \quad (11)$$

其中  $T_m = (t_1, t_2, \dots, t_m)$ ，得到标准化因变量预测值

$$\hat{F}_0 = T_m B_m \quad (12)$$

最后将预测值还原到原始数据尺度。

#### ⑤ 模型评估与应用

通过交叉验证等方法评估模型预测性能，利用构建好的模型进行预测或分析变量间关系。

### (二) 数据来源与变量说明

#### 1. 人口吸引力

根据百度智能云、百度商业智能实验室发布的《蝶变中的未来之城——雄安新区 2021 年大数据研究报告》<sup>[11]</sup>，人口吸引力的计算公式为：

$$\text{人口吸引力} = \frac{\text{当月区域内新增常住人口量}}{\text{区域常住人口总量}} \quad (13)$$

其中当月区域内新增常住人口量定义为三个月之前非区域常住人口，但是在当月成为区域常住人口的居民数量。

#### 2. 多维度评价体系

本研究主要致力于雄安新区人口吸引力的影响因素，为保证数据的真实性，数据主要来源为百度智能云、百度商业智能实验室发布的《蝶变中的未来之城——雄安新区 2021 年大数据研究报告》和《未来之城 雄姿初显——雄安新区 2022 年大数据研究报告》<sup>[12]</sup>、国家统计局、河北统计年鉴以及中国雄安官网等权威平台并进行整合。

研究聚焦于剖析雄安新区人口吸引力的影响因素，构建了系统的因素分类体系，所构建的评价体系严格遵循综合评价构建体系原则：

- 系统性原则：从经济、人才、建设、生活四个有效维度全面覆盖城市发展要素，呼应李国平（2018）<sup>[3]</sup>提出的“雄安新区四大战略”；
- 科学性原则：引入 POI（兴趣点）、网络声量趋势等客观数据，结合政策进

行分析（如“雄才计划”）提升指标效度；

- 动态性原则：选取“雄安站周围 POI 总量变化趋势”“高校迁入声量趋势”等时序指标，反映新区建设的变化情况和阶段性特征；
- 导向性原则：以“人口吸引力”为核心目标，指标选择聚焦可干预变量（如政策热度、基建投入），为政策优化提供依据。

具体内容如表 1 所示：

表 1 多维评价体系		
维度	指标	具体因素
经济因子	经济活力	科技创新类企业数量
	就业机会	企业数量 央企迁入网络声量趋势
人才因子	人才政策	“雄才计划”搜索关注度 “雄才计划”网络声量趋势
	教育资源	高校迁入声量趋势
建设因子	基础设施	雄安站周围 POI 总量变化趋势
	城市配套	路网总里程变化趋势
生活因子	生活品质	生活品质指数
		休闲广场 POI 变化
		公园 POI 变化
	环境感知	空气质量 民众对于环保的关注度

（三）缺失值处理

鉴于雄安新区正处于快速发展的初期阶段，统计体系尚在逐步完善与健全的过程中，部分细分领域的数据统计工作可能存在不够成熟的情况。同时，不同时间段政府所颁布的政策存在差异，这在一定程度上影响了数据的收集与整理，导致所收集的数据存在缺失值现象。例如，科技创新类企业数量、企业数量以及雄安站周围 POI 总量变化趋势等数据，均存在部分月份数据缺失的问题。

为保证研究严谨性，针对这些缺失值，在因子分析和预测阶段选取了可供研究的部分数据，从而只出现了少量缺失值。针对剩余少量缺失值，本研究采用回归插补法进行了处理。

表 2 部分数据展示

时间	吸引力	科技创新类企业数量	企业数量	雄安站周围 POI 总量变化趋势	央企迁入网络声量趋势
2021.11	0.62	3704	34567	0.968	0.17
2021.12	0.84	3753	35182	0.969	0.57
2022.01	0.94	3797	35500	0.972	0.41

回归插补法的核心是利用变量相关性构建预测模型。例如，处理科技创新类企业数量等缺失值时，将其设为因变量，选取企业数量、“雄才计划”搜索关注度等关联变量作为自变量。基于完整数据样本训练线性回归模型：

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \varepsilon \quad (14)$$

其中， $y$  是因变量， $x_1, x_2, \cdots, x_n$  是自变量， $\beta_0, \beta_1, \beta_2, \cdots, \beta_n$  是模型的参数， $\varepsilon$  是误差项，表示实际值与模型预测值之间的差异。运用最小二乘法优化参数，让模型精准把握变量间线性关系。训练完成后，把缺失值样本的自变量代入模型，预测出缺失的因变量值，完成缺失值填补。从而科学降低缺失值对研究结果的影响，保障数据完整连续，为后续分析筑牢数据基础，提升研究结论的可信度。

#### （四）最大值归一化

作为一种关键的数据预处理技术，在诸多研究领域发挥着重要作用。它旨在将数据集中的原始数据值，通过特定的数学变换，统一映射到一个标准的取值区间，以此来消除量纲的影响。本研究中使用归一化时将数据放缩到[0,1]区间。其核心原理基于数据集中的最大值，具体计算方式为：对于任意一个数据集，设其中的最大值为 $x_{\max}$ ，对于数据集中的每一个数据点 $x$ ，经过最大值归一化处理后的结果 $x_{\text{norm}}$ 由公式得出：

$$x_{\text{norm}} = \frac{x}{x_{\max}} \quad (15)$$

通过最大值归一化处理，能够使不同类型的数据在同一尺度下进行分析，提升研究结果的准确性与可靠性。

### 三、雄安新区发展现状概述

#### （一）经济因子

经济因子主要分为经济活力、就业机会两个指标，其中，经济活力涵盖科技创新类企业数量、企业数量、央企迁入网络声量趋势三个因素。

##### 1. 经济活力

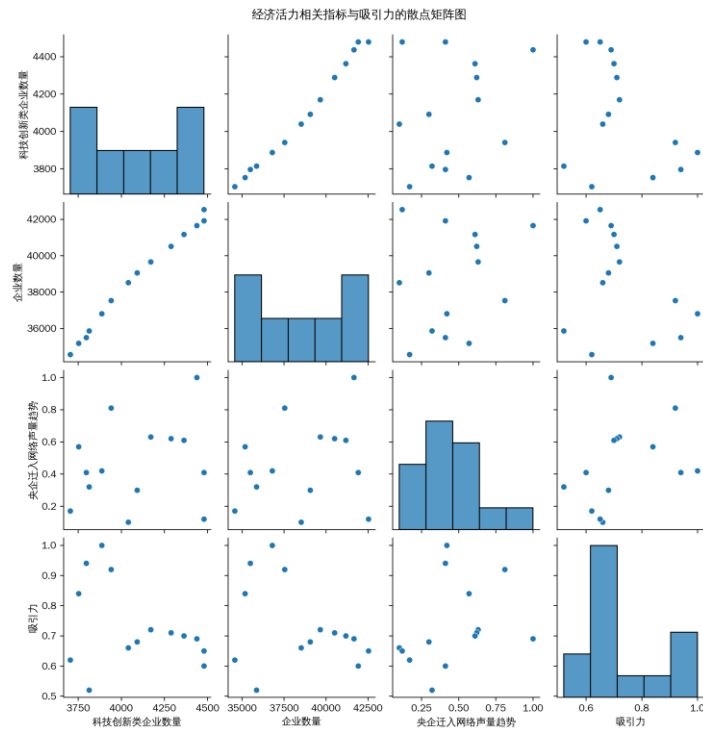


图 6 经济活力相关指标与吸引力散点矩阵图

针对**科技创新类企业数量**的分析，如图 6 显示：通过直方图可见，该指标数据集中于 3800–4200 家区间，体现出显著的集中趋势，且离散程度相对较低。散点图分析表明，科技创新类企业数量与吸引力呈弱正相关关系，随着企业数量增加，吸引力呈现轻微上升倾向，表明科技创新类企业数量增长对吸引力具有一定正向影响，但其线性关联强度较弱。

##### 2. 就业机会

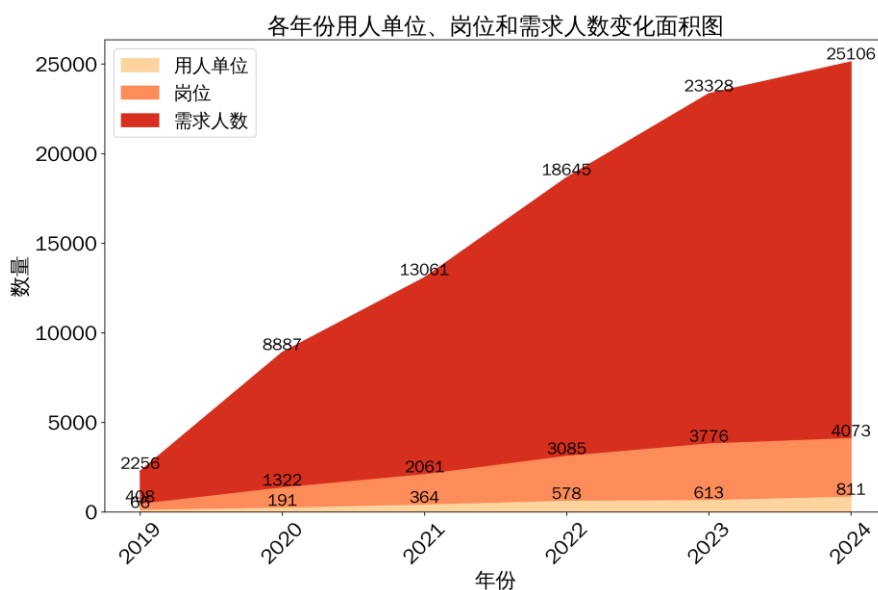


图 7 2019-2024 年就业相关数据变化面积图

由图 7 观察到：从**需求人数**上来看，2019 年需求人数约 1848 人，2024 年增至 21033 人，需求人数持续攀升。这表明就业市场对劳动力的需求总量不断扩大，释放出“岗位缺口大”的信号，直接吸引外部人口流入就业。

在**岗位与用人单位**角度上看，岗位数量从 2019 年的 342 个增至 2024 年的 3262 个，增幅 8.5 倍，用人单位数量从 66 家增至 811 家。就业主体（用人单位）与岗位供给的双重增长，意味着就业选择更丰富，能满足不同技能、背景人口的就业需求，进一步增强区域对人口的吸引力。

## （二）教育因子

这一部分聚焦于“雄才计划”搜索关注度、“雄才计划”网络声量趋势、高校迁入声量趋势三个因素，将其归类于人才政策、教育资源两个维度展开剖析。其中，“雄才计划”相关因素体现人才政策的影响力与传播态势，高校迁入声量趋势则从教育资源角度，反映区域在教育领域的动态变化及其对人口吸引力的潜在影响。

### 1. 人才政策

图 8 为人才政策相关因素与吸引力的关系可视化结果：



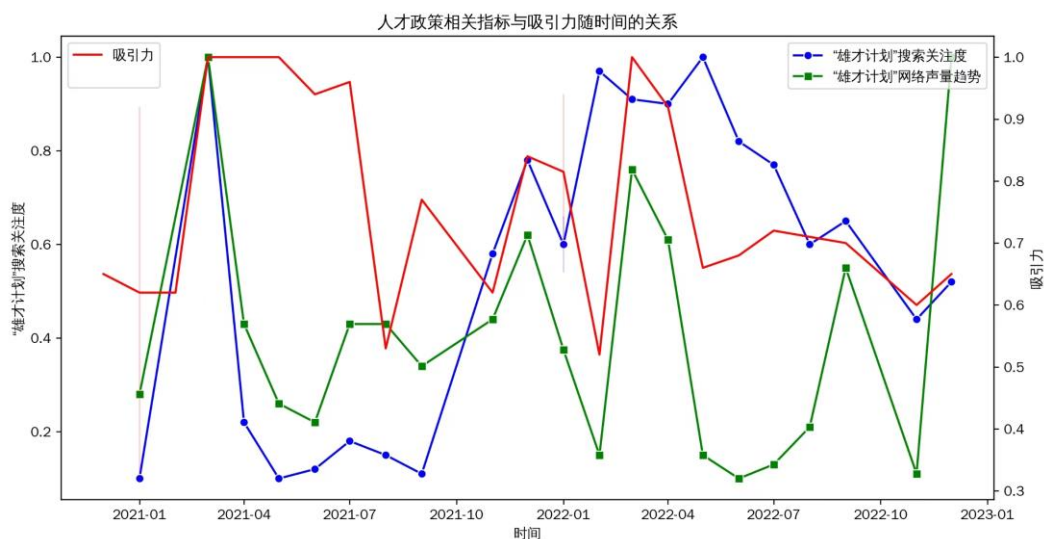


图 8 人才政策相关因素与吸引力随时间关系折线图

由图 8 直观观察到：2021 年 3 月 “雄才计划” 搜索关注度、网络声量与区域吸引力同步达峰，验证政策热度对吸引力的直接推动作用。2021 年末-2022 年初，搜索关注度上升与吸引力增长趋势一致，表明该指标提升对吸引力有促进作用。2022 年初网络声量高峰与吸引力高位同步，体现二者存在关联；但 2022 年末网络声量上升时吸引力未提升，说明网络声量与吸引力并非简单线性关系，区域吸引力受多元因素综合影响。

## 2. 教育资源

图 9 为高校迁入声量趋势与吸引力关系可视化结果：

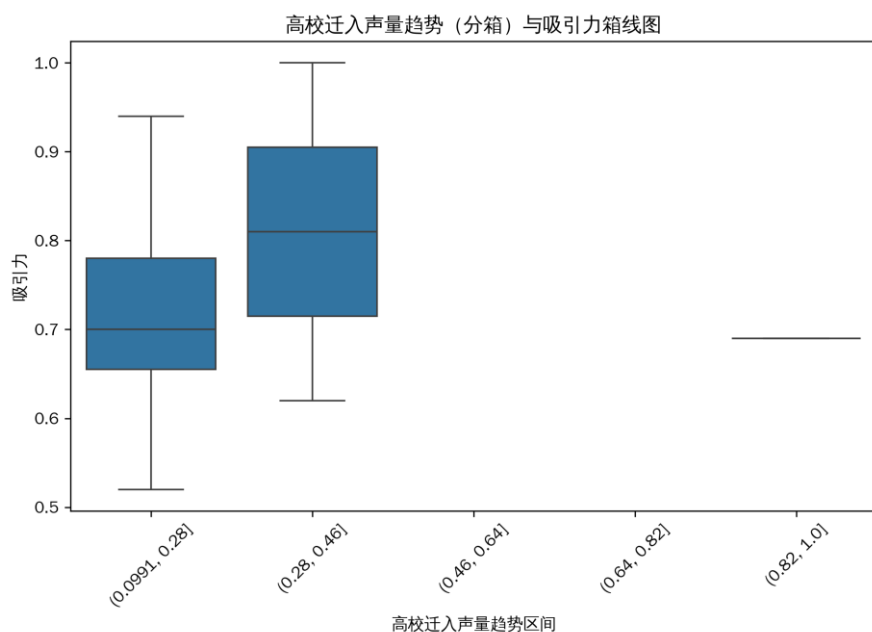


图 9 高校迁入声量趋势（分箱）与吸引力箱线图

低声量区间（0.099, 0.28]：吸引力中位数 0.7，四分位数间距（Inter - Quartile Range, IQR）覆盖 0.7-0.8（50% 数据集中于 0.7+），须线延伸至 0.5-0.9+，离散度较低但存在极值。

中声量区间（0.28, 0.46]：中位数提升至 0.8，IQR 扩展至 0.7-0.9（整体水平显著提升），须线覆盖 0.6-1.0，离散度显著增大。

随着高校迁入声量从低区间向中区间提升，吸引力中位数和 IQR 均呈现系统性上移，可知二者存在正相关关系。

### （三）建设因子

如图 10 所示，是雄安新区各类POI数量变化对比的棒棒糖图：

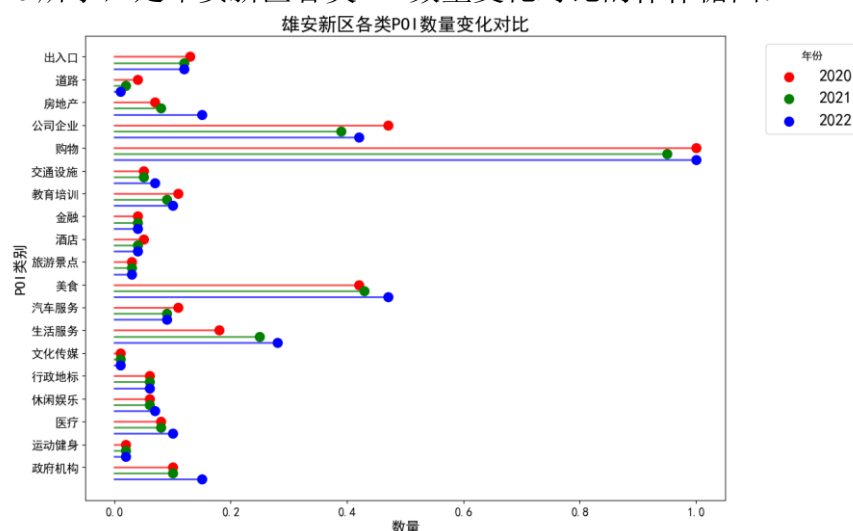


图 10 雄安新区各类 POI 数量变化对比棒棒糖图

基于雄安新区各类 POI 数量变化对比分析，其时空演进呈显著差异化特征：

从空间维度趋势上看，雄安新区各类 POI 数量占比最多的是购物、公司企业和美食三类，生活服务、房地产类增长较为迅猛；交通设施、美食、休闲娱乐、医疗、政府机构等 POI 数量呈现略微增长趋势。

从时间维度趋势上看，2020–2022 年多数类别 POI 数量递增，映射雄安新区建设中基础设施与服务功能的渐进完善。如交通设施、生活服务类 POI 在 2021–2022 年增速跃升，直观体现城市配套优化。

### （四）生活因子

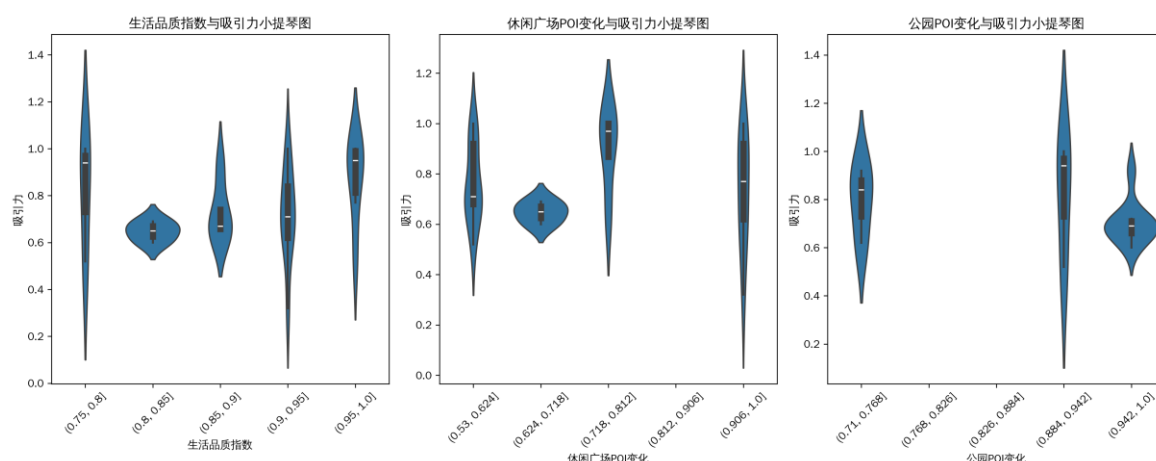


图 11 生活品质相关因素与吸引力关系小提琴图

为了清晰阐述，对图 11 包含的三类指标分别进行深入分析探究：

### 1. 生活品质指数

低数值区间(0.75, 0.8]：吸引力集中于 0.6–0.8，分布较散。

高数值区间(0.95, 1.0]：吸引力向 0.8–1.2 高位聚集，密度更高。

这一分布特征验证了生活品质指数与吸引力的正向相关性，表明生活品质指数提升对吸引力具有促进作用，高品质生活环境更易形成人群集聚效应。

### 2. 休闲广场 POI 变化

图中呈现整体效果并不明朗，但休闲广场 POI 在[0.624,0.718]至[0.718,0.812]区间吸引力呈明显跃升现象，公共休闲空间的完善对区域吸引力具有正向推动作用，反映出城市公共服务设施优化对人群吸引力的提升效应。

### 3. 公园 POI 变化

低数值区间(0.71, 0.768]：吸引力集中于 0.6–0.8。

高数值区间(0.884, 0.942]：吸引力向 0.8–1.2 高度聚集，密度大幅升高。

这一现象证明公园 POI 变化幅度与吸引力呈正向关联，公园设施优化对吸引力有提升作用。

综上，三类变量基本呈现“数值区间越高，吸引力高位分布越集中”的规律，验证生活品质指数提升、休闲广场与公园 POI 变化幅度增大，均对吸引力有正向促进作用。优化生活品质相关要素（如提升生活品质指数、完善休闲与公园设施），是增强区域吸引力的有效途径。

#### 四、人口吸引力多维度因素分类分析

本研究采用灰色关联分析（Grey Relational Analysis, GRA）探究各维度因子与人口吸引力的关联程度。灰色关联分析适用于小样本、贫信息系统的关联性评价<sup>[16]</sup>，能够有效处理雄安新区建设初期数据量有限的问题。

通过灰色关联分析，得到各维度整体关联度：

表 3 灰色关联分析关联度		
维度	关联度 ( $\Gamma$ )	排序
经济	0.814	1
人才	0.746	2
生活	0.737	3
建设	0.660	4

表 3 中，经济维度关联度最高（ $\Gamma=0.814$ ），表明产业发展是人口吸引力的核心驱动因素；建设维度关联度最低（ $\Gamma=0.660$ ），说明基础设施建设对人口吸引力的提升作用尚未充分显现。

具体地，立足于各个维度，对其中具体因素的关联度指数进行详细探究并分析：

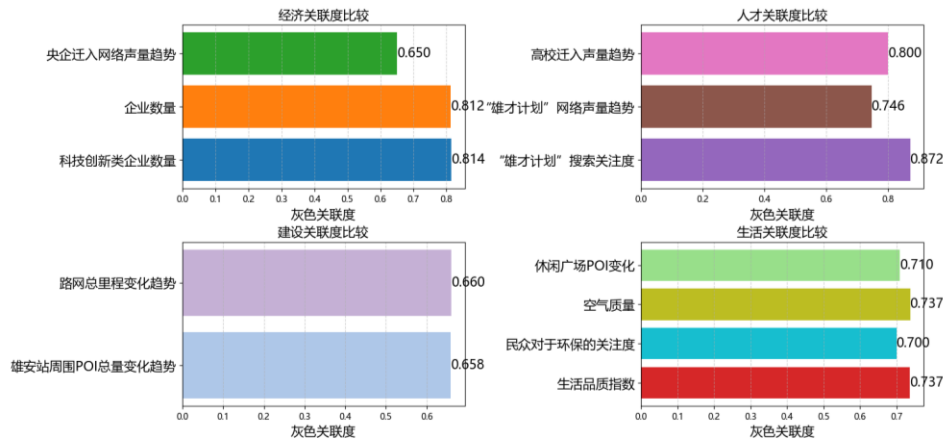


图 12 雄安新区人口吸引力-关联度比较分析

- 经济维度：**科技创新类企业数量（ $\Gamma=0.814$ ）与企业数量（ $\Gamma=0.812$ ）形成双核心驱动，而央企迁入声量贡献较低（ $\Gamma=0.650$ ）。
- 人才维度：**政策关注度（ $\Gamma=0.746$ ）略高于高校迁入（ $\Gamma=0.742$ ），表明政策宣传对人才吸引的即时效应显著。
- 生活维度：**生活品质指数（ $\Gamma=0.737$ ）是核心因子，环保关注度（ $\Gamma=0.710$ ）与空气质量（ $\Gamma=0.697$ ）呈现边际递减效应。

4. **建设维度**：POI 总量（ $\Gamma=0.660$ ）与路网里程（ $\Gamma=0.658$ ）关联度接近，表明基础设施建设需注重功能整合。

随年份具体关联系数变化趋势图如下：

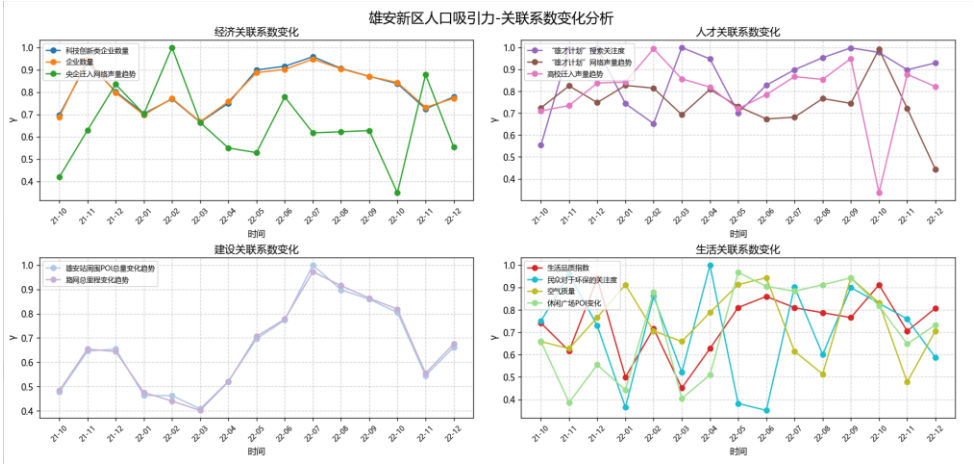


图 13 雄安新区人口吸引力-关联系数变化分析

图 13 展示了 15 个月间雄安新区人口吸引力与四大维度（经济、人才、建设、生活）的灰色关联度变化。整体呈现波动上升趋势，关联度均值从 0.65（2021.10）提升至 0.82（2022.12），表明新区发展对人口的综合吸引力持续增强。

## 五、雄安新区人口吸引力的预测分析

### （一）Arima 时间序列预测模型

雄安新区建设发展期间，时间本身具有较强趋势性。时间序列分析中的 ARIMA 模型，在处理具有趋势性和季节性的数据方面表现出色，能够有效挖掘数据内在规律，为人口吸引力预测提供有力支持。

依据《国务院关于河北雄安新区总体规划（2018—2035 年）》，雄安新区建设各阶段目标如下图：

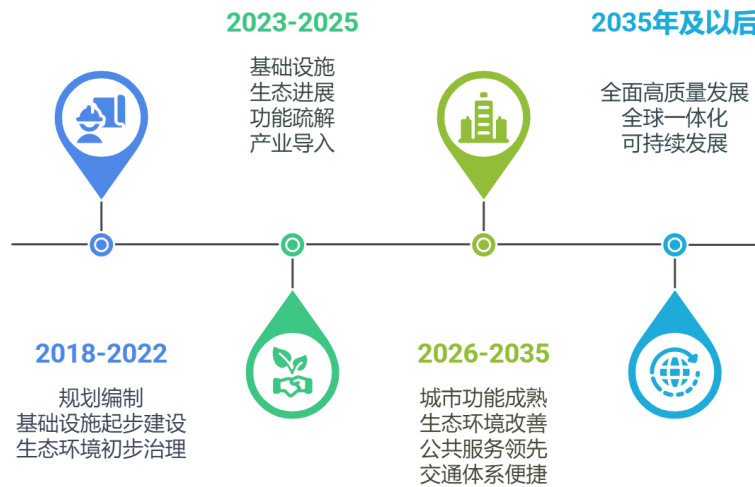


图 14 雄安新区建设各阶段目标

由于数据收集的滞后性，本文收集 2020 年 12 月至 2022 年 12 月共 25 个月的人口吸引力指数进行 Arima 时间序列预测，而鉴于模型具有普适性，对雄安新区人口吸引力指数发展趋势的预测仍具有重要借鉴意义。

#### 1. 模型参数选取

AIC（赤池信息准则）用于衡量模型的拟合优度和复杂度，AIC 值越小，模型在拟合数据和复杂度之间的平衡越好，也就是该模型更优。本研究使用 ARIMA 模型<sup>[13]</sup>的自动选择，展示在逐步搜索过程中 AIC 数值变化：

表 4 AIC 数值变化

模型	AIC 值	拟合时间（秒）
ARIMA(2,0,2)(0,0,0)[0]	-1.292	0.07
ARIMA(0,0,0)(0,0,0)[0]	60.498	0.01
ARIMA(1,0,0)(0,0,0)[0]	2.774	0.01
ARIMA(0,0,1)(0,0,0)[0]	37.357	0.01
ARIMA(1,0,2)(0,0,0)[0]	inf	0.05

ARIMA(2,0,1)(0,0,0)[0]	inf	0.06
ARIMA(3,0,2)(0,0,0)[0]	-1.482	0.07
ARIMA(3,0,1)(0,0,0)[0]	-0.937	0.04
ARIMA(4,0,2)(0,0,0)[0]	0.753	0.09
ARIMA(3,0,3)(0,0,0)[0]	0.567	0.09
ARIMA(2,0,3)(0,0,0)[0]	inf	0.07
ARIMA(4,0,1)(0,0,0)[0]	0.984	0.05
ARIMA(4,0,3)(0,0,0)[0]	inf	0.11

从表中可以看出，部分模型的 AIC 值为 inf，表明这些模型在拟合过程中存在无法收敛或数值不稳定的问题，不予选取。而(0,0,0)[0] 表示不考虑季节性成分，即该模型只关注非季节性的自回归和移动平均关系。综合比较，ARIMA (3, 0, 2)(0, 0, 0)[0] 模型的 AIC 值最小，为 - 1.482，因此被确定为本次研究的最优模型。

## 2. 残差检验

残差的正态性是评估 ARIMA 模型拟合效果的重要指标之一。

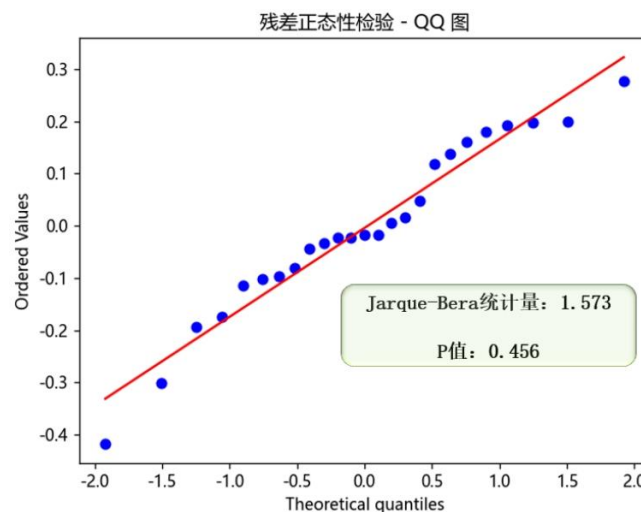


图 15 残差正态性检验 - QQ 图

由 QQ 图看出：大部分点较为接近直线，但在两端存在一定程度的偏离，表明残差近似服从正态分布。

为进一步验证，使用 D'Agostino-Pearson 正态性检验构造检验统计量进行检验，结果表明：在 5% 的显著性水平下，p 值为 0.455，显著大于 0.05，不能拒绝原假设，表明残差服从正态分布。这说明 ARIMA (3,0,2)(0,0,0)[0] 模型的残差在正态性方面表现良好，模型能够较好地捕捉数据的特征。



### 3. 预测结果

利用经过检验的 ARIMA (3,0,2) 模型，对雄安新区 2023 年 1 月至 2023 年 6 月的人口吸引力指数进行预测，结果如下图所示：

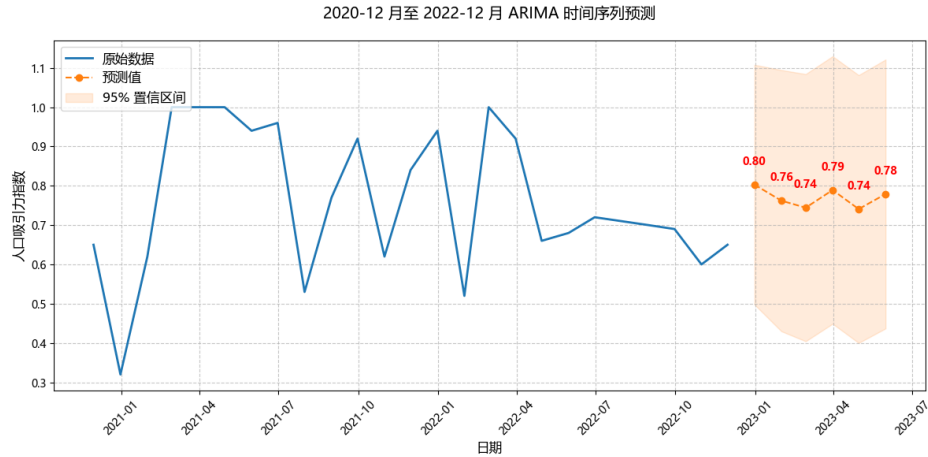


图 16 2020 - 12 月至 2022 - 12 月 ARIMA 时间序列预测

预测结果显示，在未来一段时间内，雄安新区人口吸引力指数整体呈现上升趋势，但上升过程可能存在一定的波动。具体分析：

- ① **趋势分析：**随着雄安新区建设的不断推进，各项政策的逐步落实，其人口吸引力将持续增强。这主要得益于雄安新区在经济、建设、人才、生活四方面的不断优化。
- ② **波动分析：**预测过程中出现的波动可能受到多种因素的影响，如政策调整的时间节点、重大项目的推进进度。在某些特定的季节，可能由于就业机会的变化或生活成本的波动，导致人口流动出现一定的变化；政策调整的初期，可能会引起市场的短期反应，从而影响人口吸引力。

#### （二）基于 PLSR（偏最小二乘回归）回归预测模型

本预测模型采用偏最小二乘回归法(PLSR)，该方法融合了主成分分析(PCA)、典型相关分析(CCA)和线性回归(OLS)的优势<sup>[17]</sup>，能够在本研究存在多重共线性及小样本的情况下，提取最具代表性的成分，提高模型的稳定性和预测能力。



1. 因子数量选取

表 5 潜在因子信息综合解释

潜在因子X	方差累计的 X	方差Y	方差累计的 Y	方差(R <sup>2</sup> )	调整后的 R <sup>2</sup>
1	0.236	0.236	0.268	0.268	0.211
2	0.116	0.352	0.268	0.536	0.458
3	0.044	0.396	0.226	0.762	0.697
4	0.065	0.461	0.092	0.853	0.795
5	0.049	0.51	0.019	0.872	0.802

上表展示潜在因子的信息综合解释能力，方差解释表显示：最终模型解释力达 87.2%。但考虑到变量个数同时为防止过拟合发生，调整后  $R^2=0.802$ ，优于传统灰色预测模型<sup>[7]</sup>。其中，前 4 个潜在因子可解释近 80% 的自变量信息，第 5 个因子将因变量解释率提升至 80% 以上，表明模型具有较好的预测效度。

2. 变量重要性探究

VIP 是 PLSR 模型的核心指标<sup>[14]</sup>，反映变量的重要性。VIP 值分析显示“雄才计划”网络声量(1.229) 对人口吸引力影响最大，作为关键变量。同时，VIP 值大于 1 的因素主要涵盖经济因子、人才因子、生活因子，相较于这三者，建设因子所包含因素影响较小。具体地，生活品质指数（1.180）、央企迁入声量(1.194)及科技创新类企业数量(1.022)等变量具有显著解释力。

为直观观察，可视化如下图所示：

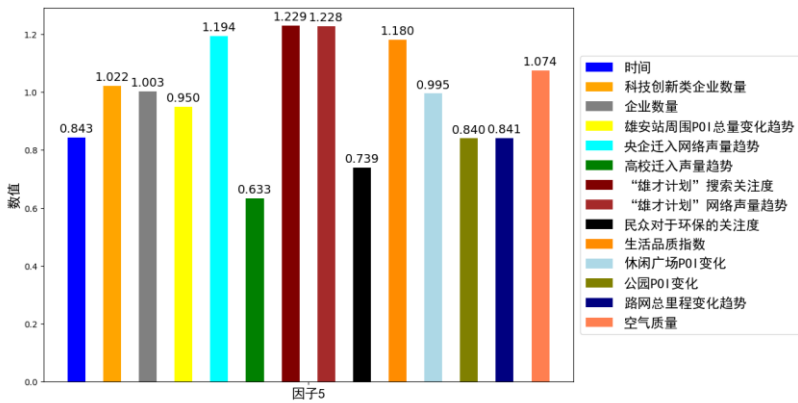


图 17 PLSR 模型的核心指标

### 3. 预测结果

如图 18，使用 PLSR 模型对 2021 年 10 月至 2022 年 12 月时间段人口吸引力指数进行拟合，预测效果良好。

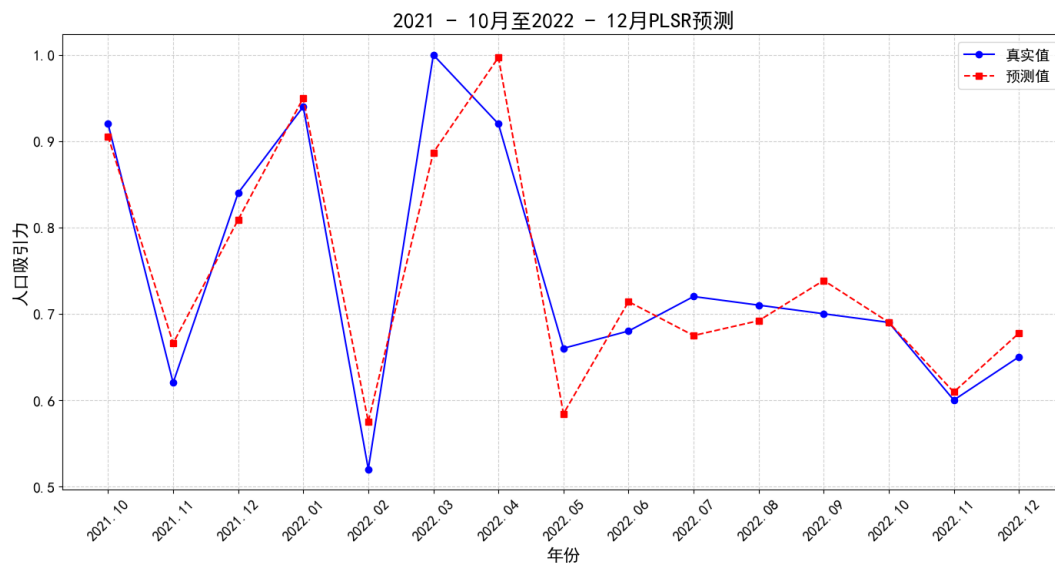


图 18 2021 - 10 月至 2022 - 12 月 PLSR 预测

## 六、研究结论

本研究通过经济、人才、建设、生活四个维度建立多维度指标体系，使用灰色关联分析、ARIMA 时间序列模型、偏最小二乘法模型三类分析预测方法，深入研究分析了雄安新区人口吸引力的影响因素，归类出正向因素并研究其影响程度，揭示了正向驱动机制与未来发展趋势。主要结论如下：

### 1. 经济驱动的核心地位

灰色关联分析结果显示，经济活力（ $\Gamma=0.814$ ）与就业机会（ $\Gamma=0.812$ ）对人口吸引力具有显著正相关关系。科技创新企业数量与企业数量的协同增长态势，印证了雄安新区通过高端产业布局形成的初步人口引力场。建设维度关联度（ $\Gamma=0.660$ ）低于预期，这与梁林（2019）<sup>[7]</sup>关于新区建设周期的研究结论一致，表明基础设施效能释放存在滞后性。

### 2. 政策与生活品质的协同效应

PLSR 模型分析得到，“雄才计划”网络声量趋势（VIP=1.229）对吸引力影响最大，这一发现支持了王艳君等（2022）<sup>[4]</sup>提出的政策宣传即时效应理论。生活品质指数（VIP=1.180）与环保关注度（ $\Gamma=0.710$ ）的显著贡献，表明“宜居+宜业”双轮驱动模式已初见成效。

### 3. 波动上升趋势

时序预测模型结果显示，ARIMA 模型基于已有数据预测的未来六个月人口吸引力指数呈波动上升趋势（ $R^2=0.802$ ）。同时，PLSR 模型拟合效果良好，可用于后续进一步研究分析。

### 4. 建设短板与需求优化

建设维度关联度（ $\Gamma=0.660$ ）在本研究的因子分析中影响最低，证明基础建设需要从规模扩张转向功能整合。以雄安站为核心的基础建设，可通过提升 POI 密度与路网效率增强服务效能，实现服务效能优化。

## 七、提升雄安新区人口吸引力的策略建议

上述研究表明，雄安新区人口吸引力呈现“经济主导（双轮驱动）、政策快响（雄才计划即时响应）、基建时滞（6-12个月）”特征。

据此，建议实施“产业升级、政策精准化、生活品质提升、基建效能革命”四大工程，构建“产业 - 人才 - 环境 - 设施”协同发展体系，推动人口吸引力持续提升。根据以上各因素表现，对提升雄安新区人口吸引力作出一些策略建议。

### （一）经济驱动：强化产业基础与就业质量

#### ① 产业升级与集群化发展

聚焦人工智能、数字经济等战略性新兴产业（ $\Gamma=0.814$ ），推行“链长制”完善产业链条，对科技创新企业（ $\Gamma=0.814$ ）实施税收减免与专项研发补贴。搭建“产学研”创新平台，推动央企、高校与科研机构共建“雄安创新联合体”，加速科技成果转化（VIP=1.022）。

#### ② 精准匹配就业需求

就业吸引人才，建立“产业-人才”动态匹配机制，定期发布重点产业人才缺口报告（ $\Gamma=0.812$ ），引导高校专业设置与产业需求对接。对于应届毕业生和技能型人才进行人才补贴创新，提供“住房补贴 + 创业基金”组合政策（VIP=1.194），降低迁移成本，吸引外来人才驻扎雄安新区。

### （二）人才引育：优化政策效能与教育资源

#### ① 政策精准化与宣传创新

对“雄才计划”（VIP=1.229）实施动态评估，建立补贴标准与技能需求的联动调整机制，重点向高技能人才和紧缺岗位倾斜。利用短视频平台、高校宣讲会等渠道延长政策热度周期，提升“搜索关注度”（ $\Gamma=0.746$ ）的持续性。

#### ② 教育资源整合与高校合作

引入北京名校、医院资源建立分校/分院，短期内提升教育医疗水平（ $\Gamma=0.742$ ）。与高校共建“定向招生 + 企业实训”培养模式，对入驻科研团体提供科研经费、住房等保障。

### （三）建设提质：基础设施协同进步

#### ① 功能整合与空间优化

以雄安站为核心构建“轨道 + 公交 + 步行”一体化枢纽，周边 1 公里布局

商业、办公、居住等多功能 POI ( $\Gamma=0.660$ )，实现“15 分钟生活圈”。推进智慧交通网络建设，打造特色通勤体系。

## ② 差异化功能分区

注重居民需求，优化建设结构。打造商务区、宜居区、生态旅游区等功能板块，强化 POI 密度与城市服务效能的空间协同<sup>[15]</sup>。

# （四）生活提质：品质升级与生态价值转化

## ① 健康住宅与公共服务

推广“健康住宅”标准，新建商品房配备新风系统、智能家居设备。每 2—3 个小区共享 1 个社区服务中心，配置菜鸟驿站、社区卫生站、老年活动室等高频需求设施，减少重复建设。

## ② 生态价值转化

建设零碳社区、绿色产业园，巩固空气质量的边际效益。选择容西安置区开展光伏屋顶、垃圾分类回收等可复制的低成本措施，逐步构建宜居宜业的绿色新城区，避免盲目投入高技术设备。

## 参考文献

- [1] 杨震,荣玥芳,田林,等.京津冀城市网络协同发展分析及雄安新区人口规模研究[J].干旱区资源与环境,2019,33(12):8-15.DOI:10.13448/j.cnki.jalre.2019.339.
- [2] 刘洁,王亚,苏杨.城市群高质量发展背景下人口聚散规律和驱动因素研究[J].河海大学学报(哲学社会科学版),2023,25(06):103-119.
- [3] 李国平,宋昌耀.雄安新区高质量发展的战略选择[J].改革,2018,(04):47-56.
- [4] 王艳君,刘清滢,司丽丽,等.雄安新区未来人口结构及促进其高质量发展的建议[J].科技导报,2022,40(22):78-87.
- [5] 孙明, 吴敏, 等. DB1331/T 064-2023 雄安新区农产品质量安全追溯体系技术要求 [S]. 河北雄安新区: 河北雄安新区管理委员会, 2023.
- [6] 梁林,武晓洁.治理群簇视阈下国家级新区人才流动预测及治理研究——以雄安新区为例[J].科技管理研究,2020,40(17):88-96.
- [7] 梁林,曾建丽,刘兵.雄安新区未来人口趋势预测及政策建议[J].当代经济管理,2019,41(07):59-67.DOI:10.13253/j.cnki.ddjgl.2019.07.009.
- [8] 邓聚龙.灰色控制系统[J].华中工学院学报,1982,(03):9-18.DOI:10.13245/j.hust.1982.03.002.
- [9] 吕浩,吕志.ARIMA 时序模型在广东省经济动态分析中的创新探索[J].特区经济,2025,(03):119-122.
- [10] 吴拥政.农村居民消费结构变动的影响因素分析——基于新兴古典经济学与偏最小二乘回归的经验证据[J].经济经纬,2008,(06):5-8.DOI:10.15931/j.cnki.1006-1096.2008.06.022.
- [11] 百度智能云,百度商业智能实验室.蝶变中的未来之城——雄安新区 2021 年大数据研究报告 [R].2022.
- [12] 百度智能云,百度商业智能实验室.未来之城 雄姿初显——雄安新区 2022 年大数据研究报告 [R].2023.
- [13] 孙建波.基于 ARIMA 线性时间序列的埋地输油管道剩余寿命预测[J].辽宁化工,2024,53(03):471-475.DOI:10.14029/j.cnki.issn1004-0935.2024.03.030.
- [14] 陈镜宇,李卫东,武冰玉.我国经济增长的收敛程度及其影响因素——基于偏最小二乘回归方法[J].经济问题探索,2024,(11):21-36.

- [15]李江,吴玉鸣.空间计量经济学前沿理论、方法与应用研究综述[J].当代经济管  
理,2024,46(06):30-41.DOI:10.13253/j.cnki.ddjjgl.2024.06.004.
- [16]Tripathy S ,Tripathy K D .Multi-response optimization of machining process para  
meters for powder mixed electro-discharge machining of H-11 die steel using grey  
relational analysis and topsis[J].Machining Science and Technology,2017,21(3):3  
62-384.
- [17]K M K ,Patrick G ,Youngseung L .Identification of Drivers of Liking for Bar-Typ  
e Snacks Based on Individual Consumer Preference.[J].Journal of food science,20  
16,81(1):S174-81.

## 附录

为确保研究的准确性与一致性，对关键名词进行定义，如下表所示：

类别	名词	定义
城市 产业 分析	POI	Point of Interest，具有特定地理坐标位置，且具有一定地理意义或商业价值的地理实体或地点。
	功能区 域	基于机器学习方法，根据一个区域的 POI 分布所推算出的该区域的主要功能。本文将功能区域分为 6 类：消费、金融、企业、政府、景区、其他。
	人口吸 引力	每个月的人口吸引力定义为当月区域内新增常住人口量比上区域内常住人口总量。其中当月区域新增常住人口量定义为三个月之前不是区域常住人口，但是在当月成为区域常住人口的居民数量。
	动态人 口密度	过去一年每月独立移动设备数除以城市总面积。
网络 舆情 分析	关注度	关注度是指对各统计指标的搜索总频次，然后再使用最大值归一化方法将整个数据缩放到 [0,1] 区间的结果值。
	声量趋 势	声量趋势是指每个统计指标产生的网络话题总体的传播量，包括阅读、转发、点赞、评论等，并使用归一化方法将整个数据放缩到 [0,1] 区间的结果值。