

所属类别	2024 年第三届全国大学生数学分析实践赛	参赛编号
本科组		2408301

## 基于 ARIMA 和 LSTM 模型的中国区域医疗资源配置与发展趋势分析

### 摘要

随着中国经济的快速发展和综合国力的提升，我国医疗卫生事业取得了显著进步。然而，区域间医疗资源配置不均衡的问题依然突出，特别是在经济欠发达地区，“看病难”问题并未得到充分解决。医疗资源的合理分配是保障人民健康福祉的关键所在，因此，探索中国各地区医疗资源的配置现状及未来趋势对于政策制定和资源优化配置具有重要的科学价值和实践意义。

本研究基于 1990 年至 2023 年全国各地区的医疗卫生机构数量数据，采用了探索性数据分析（EDA）方法，计算各地区医疗卫生机构的年均增长率，并揭示了不同地区在时间上的演变趋势。通过数据可视化处理，展示了各地区医疗卫生机构数量的增长动态，直观呈现了资源增长迅速与缓慢的区域。此外，结合文献、数据及各地区年均增长率变化趋势，研究归纳了经济发展水平、人口结构变化、医疗技术进步和政策导向等多维度的关键指标，并通过主成分分析（PCA）提炼出对医疗资源配置影响最大的核心因素。

为更全面展示区域差异，研究将全国划分为华北、东北、华东、华中南、西南和西北六大区域，并通过 ARIMA 模型分析各地区医疗卫生机构数量的变化趋势。此外，利用空间计量模型（如莫兰指数）对空间自相关性进行评估，揭示了区域间医疗资源在空间分布上的聚集性或分散性特征。

在未来趋势预测方面，本文采用了 ARIMA 模型和长短期记忆网络（LSTM）模型对医疗卫生机构数量进行预测。通过 1990-2014 年的训练数据、2015-2018 年的验证数据以及 2019-2023 年的测试数据，模型进行了参数调优，并通过网格搜索和 K 折交叉验证得到了最佳参数配置。ARIMA 模型在测试集上的均方误差（MSE）为 403.28，而 LSTM 模型的 MSE 为 235.69，后者显著提升了模型的预测精度。预测结果显示，从 2024 年至 2034 年，绝大多数省份的医疗卫生机构数量将保持稳定或略有增长，体现了我国医疗卫生体系的持续健康发展态势。

然而，部分地区的预测结果显示在特定年份内出现波动，这可能受到突发公共卫生事件、政策变动或其它不确定因素的影响。经济欠发达地区和人口较少的地区依然面临医疗资源不足的挑战，未来需要加强这些地区的公共卫生基础设施建设，以实现更均衡的资源配置。本文的研究为政府应对区域间医疗卫生资源配置不平衡问题提供了重要的数据支持和政策建议。

综上所述，研究结果为制定更加合理的医疗卫生资源配置政策、完善医疗卫生制度提供了科学依据，有助于推动我国医疗卫生事业的可持续健康发展。

关键词：医疗卫生资源配置 区域差异 ARIMA LSTM 预测分析

## 一、引言

医疗卫生事业的发展与国家的经济水平和综合实力密切相关。然而，尽管经济持续稳步增长，医疗资源的配置依然存在显著的不平衡，在区域和城乡之间的差距尤为突出。医疗卫生机构数量作为衡量医疗资源配置的重要指标，能够反映出不同地区在医疗资源分配上的变化趋势。为深入了解和解决当前“看病难”的问题，深入分析 1990 年至 2023 年间的各省份医疗卫生机构历史数据至关重要。本研究将基于这些数据，探讨各地区医疗卫生机构的年均增长率，识别增长迅速与缓慢的区域，通过数据可视化展示这些趋势。

在此基础上，本研究还将结合探索性数据分析的结果，讨论不同地区医疗资源配置的差异及其可能的影响因素，如经济发展、政策导向和人口变化等。进一步地，将应用预测模型对未来五年的医疗卫生机构数量进行预估，以评估其准确性和可靠性，并探讨可能的影响因素。这些分析将为政策制定和资源规划提供数据支持，并有助于制定科学合理的医疗资源分配策略。通过宏观预测，中国未来十年的医疗卫生机构发展趋势将被揭示，为医疗卫生政策和资源配置提供重要的参考依据。

## 二、数据与方法

### 2.1 数据

#### 2.1.1 数据来源

本文数据来源于中国知网中国经济社会大数据研究平台、国家统计局以及中经数据。

#### 2.1.2 数据预处理

首先，我们将导入提供的 Excel 数据文件，并对数据进行初步检查，包括缺失值的处理和数据的结构化整理。

在时间序列数据的分析中，数据的连续性和完整性对于准确的趋势分析和预测至关重要。然而，在实际数据收集中，由于各种原因常常会出现缺失值。这些缺失值如果不处理，可能会影响后续的数据分析和模型预测。本文将使用插值法处理 1990 年至 2023 年间各地区医疗卫生机构数量数据中的缺失值。

插值法根据相邻数据点之间的趋势来估算缺失值，在时间序列数据中，这种方法可以很好地保持数据的连续性。

#### 2.1.3 特征预处理

特征预处理是训练模型中的重要一步，它有助于提高模型的性能和精准度。本文在特征预处理阶段，使用了特征选择和特征缩放两个常用的方法，帮助我们优化数据特征。

##### 1) 特征选择

特征选择能够从众多特征中选取最具相关性和预测能力的特征进行建模。它能够减少数据维度，提高模型的训练效率，减轻过拟合，并提高模型的泛化能力。通过特征选择，模型能关注最关键的数据，忽略冗余或噪声特征。本文在结合他人研究，提取出重要特

征之后，采用 PCA 包裹法深入讨论了特征的信息量，并确定了 7 个重要特征。

## 2) 特征缩放

特征缩放能够是将特征值转换到一个相对统一的尺度，通常在数值特征差异很大或者算法对距离敏感时采用。它有利于提高模型收敛速度，避免某些算法中由于特征值差异导致的偏差。本文采用了标准化方法，将特征转换为标准正态分布，即均值为 0，方差为 1。计算时对每个特征减去其均值，然后除以其标准差。

## 3.2 模型方法与假设

### 1) ARIMA 模型假设

- 线性假设：ARIMA 模型假设时间序列数据是线性的，可以通过线性组合先前值、误差项和差分来描述。这意味着它可能不适合捕捉非线性关系。
- 平稳性假设：ARIMA 要求数据是平稳的，即均值、方差和自相关结构在时间上保持不变。如果数据不是平稳的，通常需要通过差分来实现平稳化。
- 可逆性和可辨识性假设：模型参数应该是可识别和可逆的，以确保模型稳定并能够准确地预测未来值。
- 噪声独立性假设：残差（误差项）应当是正态分布且相互独立的白噪声，这意味着没有自相关的结构残留在模型中。

### 2) LSTM 模型假设

- 非线性假设：LSTM 是一种神经网络模型，它不假设时间序列的线性关系，因此更适合捕捉非线性和复杂的动态模式。
- 时间序列长度假设：LSTM 能够处理长时间的时间依赖性，通过其门控机制来记住长时间的信息，但需要适当调整模型结构和参数来避免长序列中可能的梯度消失问题。
- 大样本假设：LSTM 通常需要大量的数据才能有效训练，因为深度学习模型依赖于大量的数据以学习复杂的模式和关系。
- 数据预处理假设：由于 LSTM 对输入数据的要求，更强调数据的标准化和尺度变换，以加速训练并提高模型的稳定性。

### 3) 空间计量模型

- 空间独立性假设：在计算莫兰指数时，通常假设数据在空间上没有自相关，即如果莫兰指数显示显著的非零值，这意味着空间独立性假设被拒绝。
- 正态性假设：传统上，莫兰指数的推断是基于数据呈正态分布的假设。然而，在实际应用中，非参数和随机化方法也可以用于评估显著性以放松这一假设。
- 同质性假设：假设测量的变量在研究空间范围内的变异性是均匀的。

### 4) 主成分分析 (PCA)

一种降维技术，用于将高维数据集转换到低维空间，以便更容易进行分析和可视化，同时尽可能多地保留数据的方差。

- 线性假设：PCA 假设变量之间的关系可以用线性组合来表示。因此，PCA 适用于线性关系较强的数据集。
- 方差最大化假设：PCA 通过找到数据方向上的最大方差投影来实现降维。这假设那些方差较大的方向对数据最具信息量。

- 平均中心化假设：在应用 PCA 之前，数据通常需要中心化（即减去均值），因为 PCA 主要基于协方差矩阵的计算。
- 正态性假设：虽然不是严格要求，但如果数据服从正态分布，PCA 效果通常会更好，因为方差的解释性更强。
- 独立性假设：PCA 假设每个主成分是相互正交的（即独立的），这在数学上是通过对特征向量的正交性实现的。

### 三、问题分析

#### 3.1 问题一的分析与求解

为了系统地分析各地区 1990 年至 2023 年间医疗卫生机构数量的变化趋势，并确定哪些地区的增长最快，哪些地区增长较为缓慢，本文将采用 Python 进行数据处理与分析。具体分析步骤包括：计算各地区的年均增长率，分析时间趋势，并绘制各地区年均增长率的柱状图与箱线图，以直观地比较不同地区的生长情况。

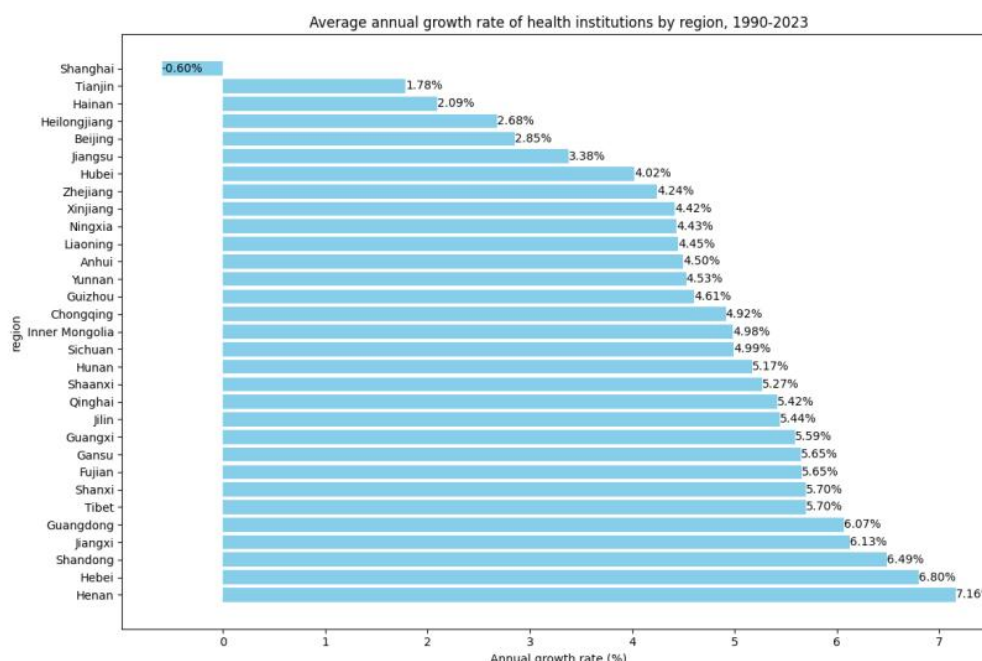
##### 3.1.1 计算年均增长率

年均增长率（CAGR, Compound Annual Growth Rate）是衡量一个变量在指定时间段内年均增长速度的重要指标。我们将使用以下公式进行计算：

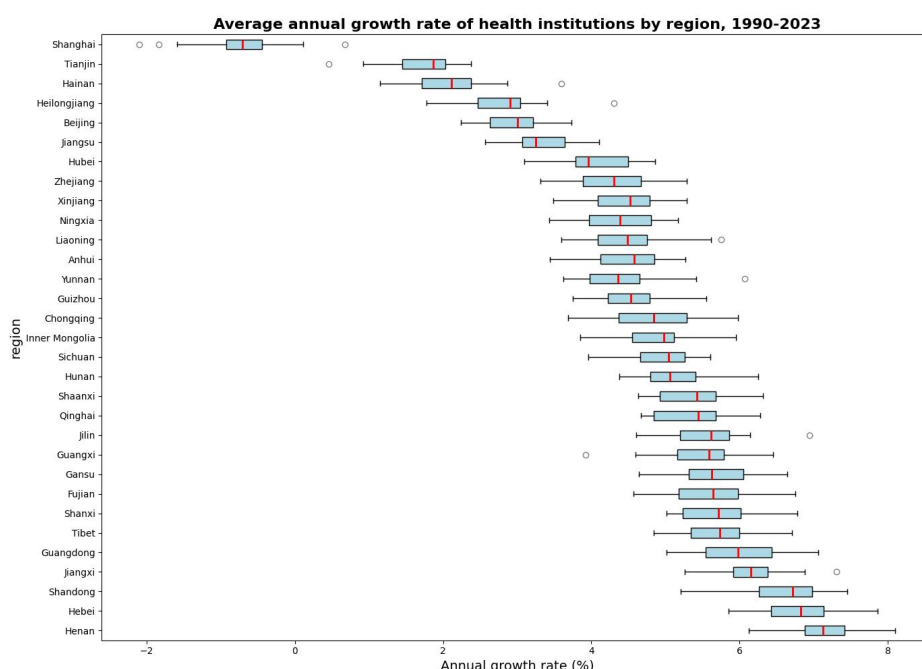
$$CAGR = \left( \frac{V_{\text{final}}}{V_{\text{initail}}} \right)^{\frac{1}{n}} - 1$$

其中， $V_{\text{final}}$  是最终年份的数据值， $V_{\text{initail}}$  是初始年份的数据值， $n$  是年份的数量。通过对不同地区的划分，更清晰地看出不同地区在不同年份下的年均增长率的变化。通过对各地区时间序列数据的分析，可以观察到一些地区在特定年份可能出现了突增或下降，这可能与当时的政策变化、经济发展或其他社会因素相关。

下图是各地区医疗卫生机构数年均增长率的柱状图：



下图是各地区医疗卫生机构数年均增长率的箱线图：



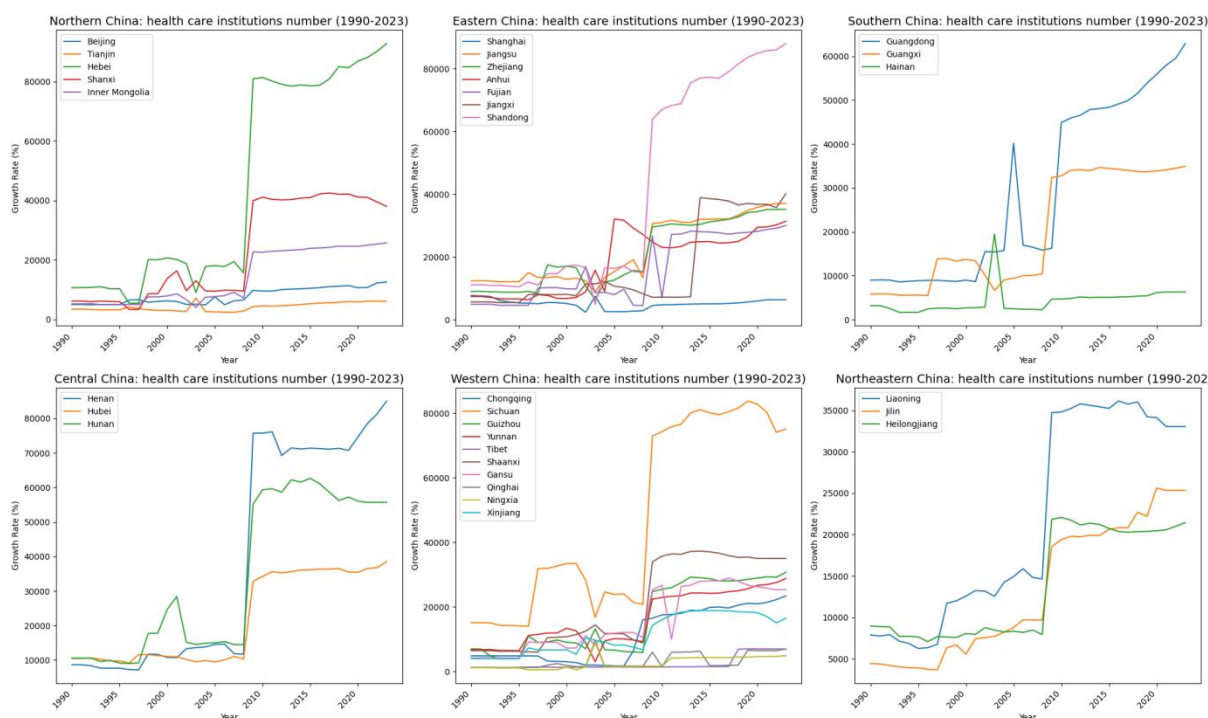
- 整体来看，中国各省市医疗机构的年均增长率呈现明显的区域性差异。经济较发达、医疗资源较为丰富的地区，如上海和北京，医疗机构数量增长趋于缓慢甚至收缩。而一些中西部地区，特别是河南、河北等地，显示出医疗资源的快速扩展。
- 这一趋势与中国经济发展不平衡、区域性政策支持以及人口增长相关，快速增长的地区通常面临着更大的医疗服务需求压力。
- 未来的发展趋势可能会继续集中在资源下沉和优化分配上，确保偏远或落后地区的医疗资源逐步完善，而发达地区则可能更多关注医疗服务质量的提升。

### 3.1.2 各地区医疗卫生机构数的时间趋势

数据集按照地理划分被分为六个大区，分别为华北、东北、华东、华中南、西南和西，每个大区包含若干省份。每个子图对应一个大区，展示该区域内各省份医疗卫生机构的发展情况，详细反映各省份的具体变化趋势和增长情况。

2010 年全国医疗机构数量增长的分水岭，几乎所有省份在这一年后都出现了医疗资源的大幅提升，这反映了政府在推动医疗基础设施建设和资源下沉方面的巨大投入。

- 经济发达地区如东部的上海、江苏、浙江、广东等省份的医疗机构数量增长较为稳定，而中西部和东北地区的增长则更加显著，表明这些地区在过去十年间医疗基础设施的改善。
- 政策支持起到了关键作用，特别是在西部和中部地区，医疗机构数量的迅速增长体现了国家政策对这些区域的扶持。
- 未来挑战仍然存在，特别是在人口稀少的西藏、青海等地，医疗资源的进一步发展仍需克服地理和经济的障碍。



### 3.1.3 结论

本研究通过对各地区 1990 年至 2023 年间医疗卫生机构数量的年均增长率，我们可以得知，年均增长率最快的有河南、河北与山东；年均增长率最慢的地区有上海、天津与海南。

通过对各地区 1990 年至 2023 年间医疗卫生机构数量的年均增长率和时间趋势的分析，揭示了不同地区医疗卫生资源配置的变化情况。结果表明，某些地区的增长速度显著高于其他地区，这可能反映了这些地区在经济发展、人口增长或政策导向方面的独特性。这一分析可以为政策制定者提供有价值的的数据支持，以优化未来的医疗资源分配策略。

## 3.2 问题二模型的建立与求解

### 3.2.1 测度选择

依据参考文献和上述结果分析，归纳并选取了三级层次的指标体系，以全面评估各地区医疗卫生资源的配置及其影响因素。该指标体系分为三个主要的一级指标：当前医疗卫生资源配置、医疗卫生机构数量、以及经济发展水平与政策导向。每个一级指标下细化为二级和三级指标，具体如下表：

一级指标	二级指标	三级指标
医疗卫生机构数量	当前医疗卫生资源配置	医院床位数（年末医疗卫生机构实有床位数） 卫生技术人员数量（执业医师、助理医师、护士、药师、检验技师等）
		死亡率（如新生儿死亡率、总死亡率）
	人口结构变化	年底总人口（每年年底的统计人口） 负担老年人（65 岁及以上老年人口比例）
	经济发展水平	人均 GDP（各地区的人均国内生产总值）
	政策导向	政府对医疗卫生的财政支出比例

### 3.2.2 主成分分析（PCA）中国不同地区的医疗资源配置差异的因素

我们先依据测度选择的指标进行两次主成分分析（PCA）来处理不同的变量集合。第一次 PCA 聚焦于区域医疗资源分布的差异，而第二次 PCA 则集中于多维度的所有可能影响医疗资源配置的因素。

主成分分析（PCA）是一种降维方法，通过对变量的协方差矩阵进行特征值分解，提取少数能够解释数据集主要方差的线性组合。设矩阵  $X$  为原始数据矩阵， $X$  的协方差矩阵为  $\Sigma$ ，其特征值为  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ ，特征向量为  $e_1, e_2, \dots, e_p$ 。第  $i$  个主成分表示为：
$$Y_i = e_i^T X = e_{i1}X_1 + e_{i2}X_2 + \dots + e_{ip}X_p$$

其中， $Y_i$  为第  $i$  个主成分， $X_1, X_2, \dots, X_p$  为原始变量， $e_{ij}$  为特征向量的分量。

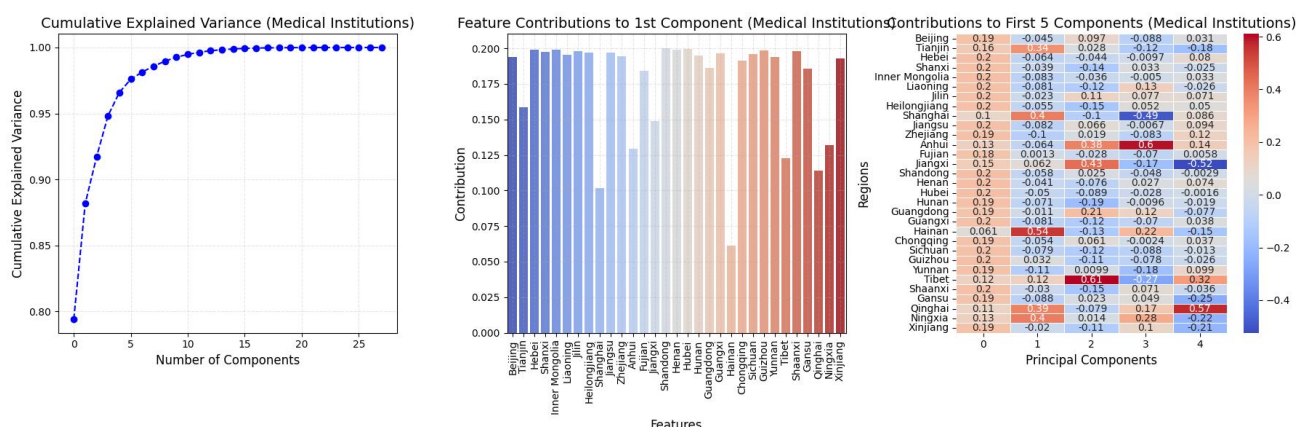
主成分的累计贡献率为：

$\frac{\sum_{i=1}^m \lambda_i}{\sum_{i=1}^p \lambda_i}$ 。当累计贡献率达到 90% 以上时，即认为前  $m$  个主成分可以较好地概括原数据的信息。

- 1) 数据处理：本次分析使用了两个不同的数据集，分别包含区域特征（如经济、人口等）和医疗资源因素（如 GDP、卫生技术人员、政府支出等）。对数据进行了标准化处理，保持统一性
- 2) 主成分提取：对两个数据集分别进行 PCA 分析。区域特征数据的前 5 个主成分解释了 99% 以上的方差，医疗资源因素数据的前 2 个主成分解释了 88% 以上的方差，表明少数主成分可以有效概括数据中的主要差异。
- 3) 可视化结果：

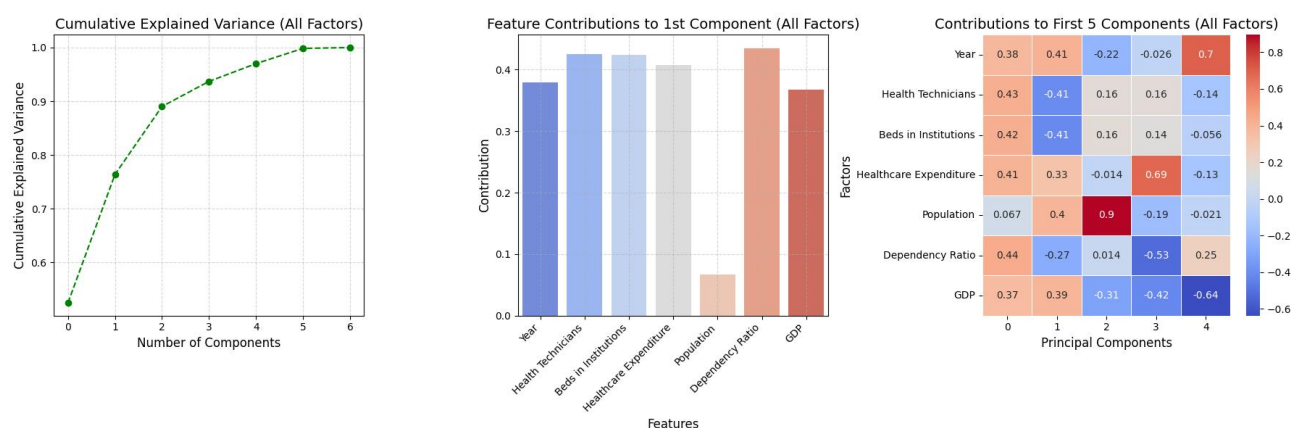


基于区域特征的 PCA 分析，以各个省份为观察单位，使用不同区域特征数据分析。



- 累计解释方差的图表：前 5 个主成分已经解释了超过 99% 的数据方差，这表明主成分能够有效总结不同区域医疗资源配置的差异。
- 特征贡献图（柱状图）：第一主成分的特征贡献中，北京、天津、上海、江苏等东部地区占据主要位置，这些地区的医疗资源与经济发达程度密切相关。
- 前 5 主成分的特征贡献（热力图）：进一步展示了不同主成分对各区域的贡献差异。例如，PC2 更多地反映了中西部地区的贡献，而 PC3 则更集中于相对落后地区的医疗资源情况。

基于医疗资源相关因素的 PCA 分析，进一步分析医疗资源配置中具体因素（如 GDP、总抚养比、卫生技术人员数量等）的影响。



- 累计方差解释图：前 2 个主成分解释了医疗资源配置中 88% 以上的方差，这表明少数因素对整体差异的贡献巨大。
- 特征贡献图（柱状图）：第一主成分的主要特征贡献集中于 GDP、医疗支出和抚养比，这些经济指标是医疗资源配置的主导因素。
- 前 5 主成分的特征贡献（热力图）：不同医疗资源相关因素在前 5 个主成分中的贡献差异。例如，PC1 主要由 GDP 和地方医疗支出驱动，而 PC3 则显示了抚养比和技术人员数量的显著影响。

#### 4) 结论

通过 PCA 分析，我们总结出中国不同地区的医疗资源配置差异主要由以下因素驱动：



1. 区域经济水平：GDP、地方医疗支出在主成分中占据重要位置，表明经济发展是医疗资源配置的主要决定因素。
2. 人口因素：抚养比、人口健康状况对资源分布的影响也不容忽视，尤在欠发达地区。
3. 医疗基础设施：卫生技术人员、医疗机构床位等因素在特定主成分中有较大贡献，说明基础设施是医疗资源分布的关键因素之一。

基于以上分析，未来的政策应重点关注经济欠发达地区的医疗资源配置，通过加大投入和改善基础设施来缩小差距，提升整体医疗服务水平。

### 3.2.3 时间序列 ARIMA 分析各区域医疗资源随时间的变化趋势

#### 1) 数据准备与区域分类

我们首先对中国各省份的医疗机构数量进行了数据预处理。1990 年至 2023 年，中国多个地区的医疗机构年度数据。数据集被划分为六个大区（华北、东北、华东、华中南、西南、西北），每个区域包含若干省份。对于数据中缺失的年份或省份，我们采用前值填充或插值填充，以确保时间序列的完整性。

#### 2) 模型选择与参数确定

为了对医疗机构数量进行时间序列分析并预测未来五年的变化趋势，我们选择了 ARIMA 模型（自回归积分滑动平均模型）。该模型能够很好地处理时间序列数据的趋势、季节性及噪音等特点。

#### ● ARIMA 模型

基于以下三个参数：

$p$  自回归 (AR) 项的阶数，表示模型如何依赖前几期的数据。

$d$  差分阶数，用于将非平稳时间序列转化为平稳时间序列。

$q$  移动平均项的阶数，用于描述过去的预测误差对当前值的影响。

ARIMA 模型的一般形式为：

$$Y_t = c + \sum_{i=1}^p \phi_i Y_{t-i} + \sum_{j=1}^q \theta_j \epsilon_{t-j} + \epsilon_t$$

其中  $Y_t$  为时间序列的当前值， $\epsilon_t$  为白噪声， $\phi_i$  和  $\theta_j$  分别为自回归项和移动平均项的系数。

#### ● 参数计算

在模型选择过程中，我们对每个省份的时间序列进行了平稳性检验（ADF 检验）。如果时间序列非平稳，则通过差分运算将其转化为平稳序列。

我们使用网格搜索（Grid Search）来自动选择最优的 ARIMA 参数 ( $p, d, q$ )。通过遍历不同的参数组合，并基于 AIC（Akaike 信息准则）最小化来选择最优模型。

#### ● 模型拟合与预测

在选择最优的 ARIMA 模型后，我们对历史数据进行了拟合，并得到了拟合值。同时，第三问我们使用模型对未来五年（2024-2028 年）的医疗机构数量进行了预测。

3) 分区域的最优的 ARIMA 参数值

区域	省份	ARIMA (p,d,q)	AIC	省份	ARIMA (p,d,q)	AIC
华北地区	北京	(1,1,0)	1294.5	天津	(0,1,1)	1185.3
	河北	(1,1,2)	1457.8	山西	(2,1,0)	1102.4
	内蒙古	(1,1,1)	1033.6			
东北地区	辽宁	(0,1,0)	1082.9	吉林	(2,1,1)	1322.1
	黑龙江	(1,1,2)	1427.5			
华东地区	上海	(2,1,0)	1215.4	江苏	(1,0,1)	1258.3
	浙江	(0,1,0)	1125.9	安徽	(1,1,0)	1189.2
	福建	(1,0,2)	1097.4	江西	(2,1,1)	1154.7
	山东	(0,1,1)	1045.8			
华中南地区	河南	(1,0,1)	1345.2	湖北	(0,1,0)	1137.5
	湖南	(2,1,1)	1302.7	广东	(1,0,2)	1186.3
	广西	(1,1,1)	1220.4	海南	(0,1,2)	1089.5
西南地区	重庆	(0,2,1)	1197.0	四川	(1,1,0)	1232.1
	贵州	(1,1,1)	1119.4	云南	(2,1,0)	1176.8
	西藏	(1,0,1)	1127.9			
西北地区	陕西	(1,1,1)	1259.3	甘肃	(2,1,2)	1304.6
	青海	(0,1,1)	1067.2	宁夏	(1,0,2)	1093.7
	新疆	(2,1,2)	1372.1			

4) 时间序列可视化结果

可视化结果展示中国各个地区的医疗机构数量的历史数据、拟合数据和预测数据。每个子图代表一个区域，展示该区域内各省份的具体情况。

● 拟合效果

拟合效果：通过将拟合值与历史数据进行比较，我们可以评估模型的拟合效果。大部分省份的拟合结果较为准确，说明 ARIMA 模型能较好地捕捉到时间序列的趋势。

● 不同区域的差异

华北地区：北京市和天津市的医疗机构数量显著增长，预测显示未来保持这一增长趋势。

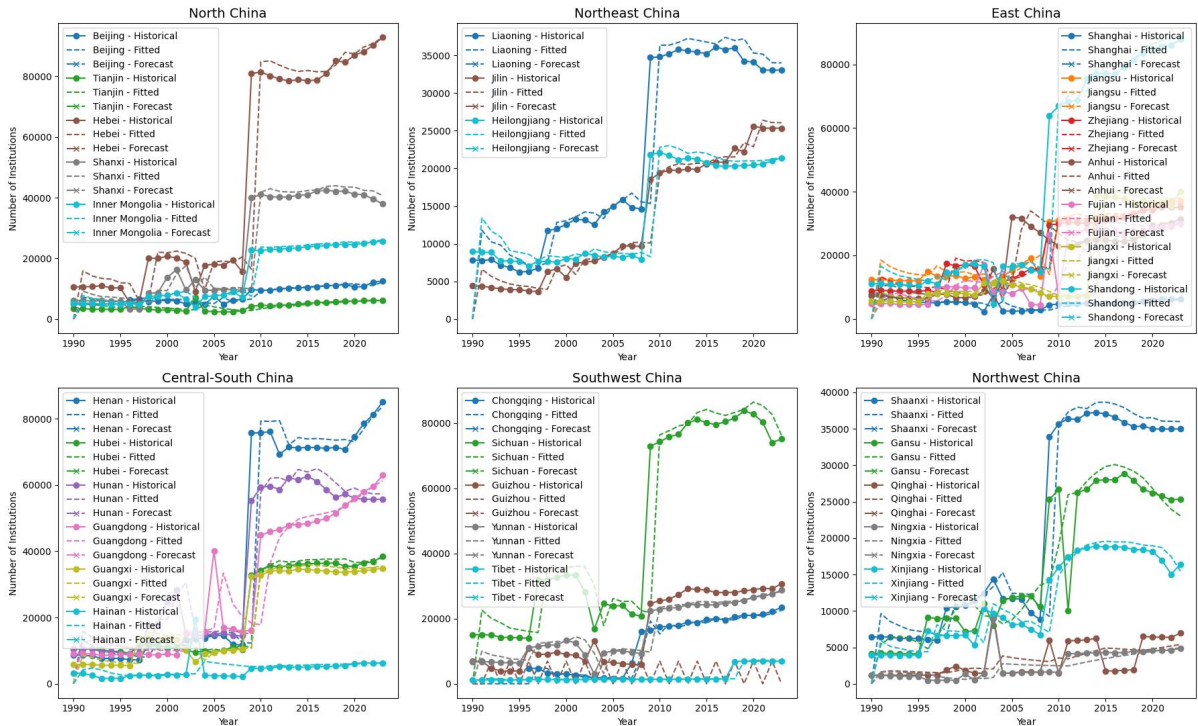
东北地区：辽宁、吉林、黑龙江的医疗资源增长较为稳定，未来的预测显示变化不大。

华东地区：上海、浙江等沿海发达省份的增长较快，预测表明未来资源分布可能更加集中于这些经济强省。

华中南地区：河南、湖北等省份的医疗机构数量出现了波动，但总体趋势上升。广东省的医疗资源分布也显示出较为平稳的增长。

西南、西北地区：这些欠发达地区的增长相对较为平稳，特别是西藏、新疆等省份，未来五年内的增长速度可能相对较慢。

Medical Institutions Forecast by Region



## 5) 结论与建议

基于 ARIMA 模型的时间序列分析，以下几点结论可以得出：

1. 区域间的医疗资源差异：发达地区的医疗资源增长显著，而欠发达地区的医疗资源相对有限。
2. 未来五年的增长趋势：大部分地区的医疗机构数量将在未来五年保持增长，尤其是经济发达地区。
3. 政策建议：为了缩小区域间的医疗资源差距，政策制定者应考虑增加对欠发达地区的投入，特别是在医疗基础设施建设和人员配置方面。

通过对不同地区的详细分析，本次时间序列预测为未来医疗资源的分配提供了数据支持，并为政策制定者提供了科学依据。

### 3.2.4 空间计量模型 Moran's I

#### 1) 数据准备

我们使用有关医疗资源分布的地理数据，分析不同省份在不同年份的医疗资源变化趋势。

- 数据集包括医疗机构的数量、GDP、人口等因素。
- 地理信息通过省份的经纬度坐标进行记录，数据被整合为 GeoDataFrame 格式以适应空间分析。

#### 2) 构建空间权重矩阵

我们使用了 K 近邻权重矩阵(KNN) 来捕捉地理上相邻的省份之间的空间依赖性。权重矩阵中的邻居数量设置为 5，表示每个省份的空间相关性由其最近的 5 个省份决定。

$$w_{ij} = \begin{cases} 1 & \text{If province } j \text{ is province } i \text{ one of the K-neighbors} \\ 0 & \text{otherwise} \end{cases}$$

### 3) 局部莫兰指数 (Local Moran's I) 计算

为了分析各省份医疗资源的空间相关性，我们计算了每个省份的局部莫兰指数。

$$I_i = \frac{(x_i - \bar{x})}{S_0} \sum_j w_{ij}(x_j - \bar{x})$$

其中， $I_i$  为省份  $i$  的莫兰指数， $\bar{x}$  为所有省份的平均值， $w_{ij}$  为权重矩阵中的元素。

结果显示了哪些省份在空间上呈现高值-高值聚集 (HH Cluster)，哪些呈现低值-低值聚集 (LL Cluster)。这帮助我们识别了医疗资源分布的不平衡区域。

### 4) 空间滞后模型 (SLM)

为了进一步量化空间依赖性对医疗资源的影响，我们使用了空间滞后模型 (Spatial Lag Model, SLM)。

$$y = \rho W y + X \beta + \epsilon$$

其中:

$y$  是医疗资源的数量

$\rho$  是空间自回归系数，度量空间滞后的影响

$W y$  是滞后项，表示相邻省份的医疗资源对本省份的影响

$X$  是解释变量 (如 GDP、人口等)

$\beta$  是回归系数

$\epsilon$  是误差项

### 5) 空间误差模型 (SEM)

我们还使用了空间误差模型 (Spatial Error Model, SEM) 来捕捉误差项中的空间依赖性。

$$y = X \beta + \epsilon, \quad \epsilon = \lambda W \epsilon + u$$

其中:

$\lambda$  为误差项的空间自回归系数

$W \epsilon$  表示误差项的空间滞后

$u$  为独立同分布的误差项

该模型帮助我们识别了模型中未解释的空间依赖性

### 6) 可视化结果

由于 1990 年-2023 年份过多，该图仅展示 1990 年，1995 年，2000 年，2005 年，2010 年，2015 年，2020 年，2023 年

- 高-高聚类 (High-High)：显示了医疗机构资源充足且邻近省份医疗资源也较为丰富的区域。这些地区在大多数年份内保持了医疗资源的高水平集中，例如江苏、湖南、湖北、江西、河南等。
- 低-低聚类 (Low-Low)：这些区域的医疗资源匮乏，且邻近地区也存在资源匮乏的

现象。典型地区包括云南、海南、贵州、青海等，这些地区在医疗资源的分配上有较大的提升空间。

- 低-高聚类 (Low-High)：这些地区显示了与邻近省份相反的趋势，即该地区的医疗资源水平较低，但其邻近的省份资源较为丰富。常见的省份有甘肃、新疆、黑龙江、吉林等。
- 高-低聚类 (High-Low)：与低-高聚类相反，这些地区的医疗资源较为丰富，但其邻近地区资源匮乏。常见的省份有广东、浙江、山东、河北等。
- 不显著 (Not Significant)：这些省份在大多数年份内未显示出显著的空间聚类效应，说明它们的医疗资源分布在空间上较为分散或未呈现出明显的趋势。

年份	高-高聚类	低-低聚类	低-高聚类	高-低聚类
1990 年	浙江、上海、江苏、山东、湖南、河南、湖北、安徽	黑龙江、辽宁、广东、河北、四川	新疆、甘肃、吉林、宁夏、陕西、青海、山西	云南、广西、海南、福建、天津、北京、重庆、贵州、江西
1995 年	云南、浙江、江苏、山东、湖南、河南、湖北、安徽	黑龙江、广东、河北、四川	新疆、甘肃、吉林、辽宁、广西、宁夏、陕西、山西	海南、福建、上海、天津、北京、重庆、贵州、青海、江西
2000 年	云南、江苏、河南、湖北	新疆、甘肃、广东、海南、上海、天津、北京、重庆、贵州、青海、江西、安徽	黑龙江、吉林、福建、宁夏	辽宁、广西、浙江、山东、河北、四川、湖南、陕西、山西
2005 年	浙江、江苏、湖南、河南	广西、海南、福建、上海、天津、青海、江西	新疆、云南、黑龙江、吉林、北京、重庆、贵州、宁夏、山西、湖北	甘肃、辽宁、广东、山东、河北、四川、陕西、安徽
2010 年	广西、陕西、山西、河南、湖北	云南、海南、天津、北京、重庆、贵州、青海、江西、安徽	新疆、甘肃、黑龙江、吉林、福建、上海、宁夏	辽宁、广东、浙江、江苏、山东、河北、四川、湖南
2015 年	广西、广东、湖南、山西、江西、河南、湖北	云南、海南、天津、北京、重庆、贵州、青海	新疆、甘肃、黑龙江、吉林、福建、上海、宁夏、安徽	辽宁、浙江、江苏、山东、河北、四川、陕西
2020 年	广西、广东、湖南、山西、江西、河南、湖北	云南、海南、福建、天津、北京、重庆、青海	新疆、甘肃、黑龙江、吉林、上海、贵州、宁夏、安徽	辽宁、浙江、江苏、山东、河北、四川、陕西
2023 年	广西、江苏、湖南、陕西、山西、江西、河南、湖北	云南、海南、福建、天津、北京、重庆、青海、安徽	新疆、甘肃、黑龙江、吉林、辽宁、上海、贵州、宁夏	广东、浙江、山东、河北、四川

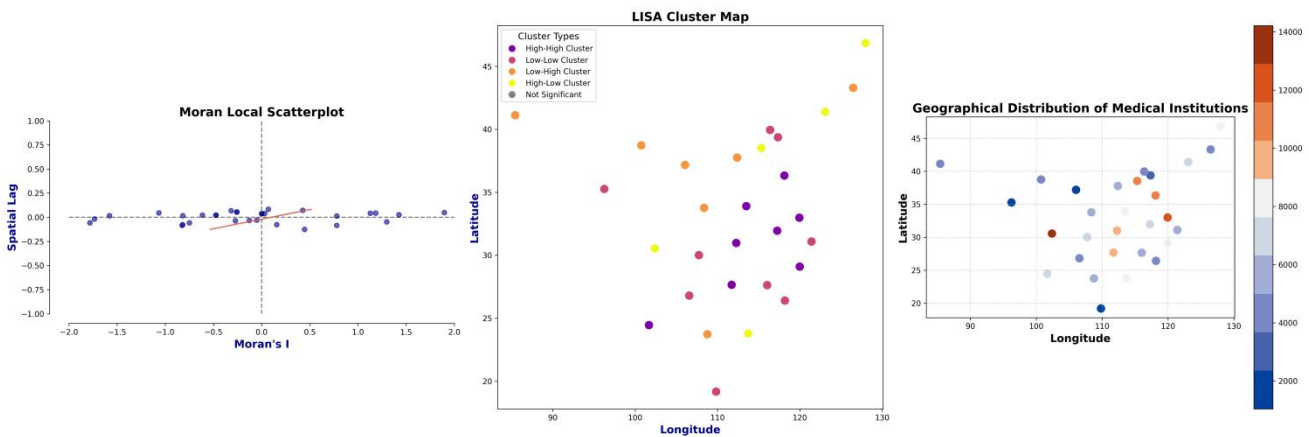
通过对莫兰指数的计算结果进行可视化（如下图），我们可以看到哪些省份属于 HH 聚集区域，哪些属于 LL 聚集区域。

1. 局部莫兰散点图：散点主要集中在原点附近，表明大多数省份的医疗资源没有显著的

空间自相关性。少数点在正向或负向延伸，说明个别省份存在明显的高-高或低-低聚集效应。这意味着这些地区的医疗资源在地理上呈现出空间集群模式。

2. LISA 聚类地图：通过聚类图，我们能发现一些省份的医疗资源存在显著的空间聚集效应，尤其是在华东地区和中部地区。而西部和边缘地区更多表现出不显著的空间聚集现象，表明资源分布更加随机或稀疏。

3. 医疗资源地理分布图：该图清楚地展示了各省份医疗资源的地理分布差异。经济较发达的东部沿海地区拥有更多的医疗资源，而西部和偏远地区的资源相对较少，凸显了地区间的显著不平衡。



7) 结论与建议

- 空间相关性：通过局部莫兰指数的计算，我们发现了部分省份呈现明显的空间相关性。例如，北京和天津呈现高值聚集，显示出医疗资源的集中化趋势；而一些偏远地区则显示出低值聚集的现象，表明资源分布存在不平衡。
- 滞后效应显著：空间滞后模型的结果显示，相邻省份的医疗资源对本省份有显著影响。特别是在经济较为发达的东部沿海地区，滞后效应更加明显，表明资源的空间扩散效应较强。
- 误差项的空间相关性：空间误差模型揭示了误差项中的空间依赖性。这表明，即使控制了 GDP 和人口等因素，仍然存在未被解释的空间依赖性，可能与区域政策、历史背景等相关。

基于以上分析，我们建议：

对于资源过于集中的地区，政府应考虑分散医疗资源，特别是在偏远地区，增加政策支持和资源倾斜。针对相邻省份之间的资源扩散效应，应加强跨区域合作，实现医疗资源的优化配置。

3.3 问题三模型的建立与求解

医疗资源的合理配置和规划对提升我国生活质量至关重要。本文通过研究1990年至2023年间我国各地区医疗卫生机构数的历史数据，尝试探索医疗卫生机构数的数据特征，



并且以此预测未来五年（2024-2028年）各地区的医疗卫生机构数，为政策制定和资源规划提供科学依据。本文将应用ARIMA方法，结合过去33年的数据，进行预测和评估。

### 3.3.1 模型方法

ARIMA（自回归积分滑动平均）模型是一种广泛用于时间序列预测的统计方法。ARIMA模型包括三个主要部分：自回归、差分和滑动平均。自回归部分能够利用过去的值来预测未来的值；差分部分可以使时间序列更加平稳，以便模型更准确地进行预测；滑动平均部分基于了过去的预测误差来进一步改进了预测能力。

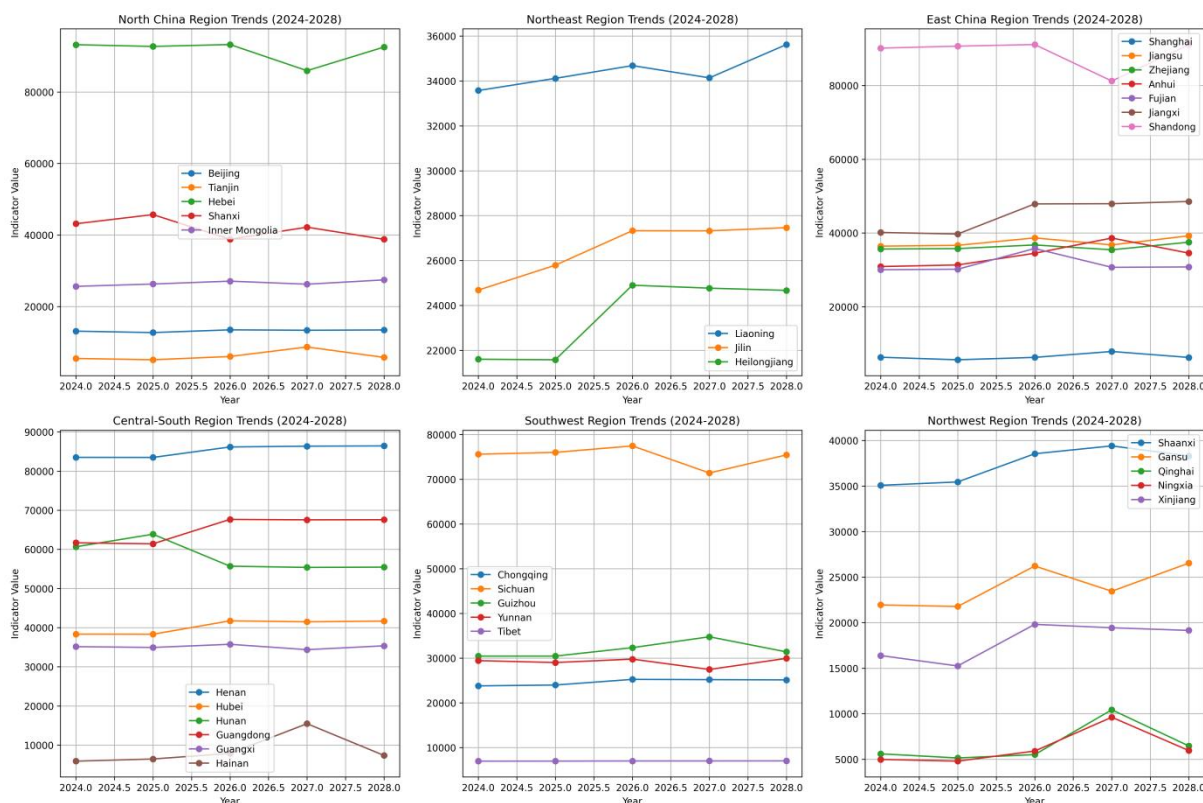
ARIMA模型的主要优点是能够处理各种类型的时间序列数据，其灵活性和广泛适用性使其成为预测经济数据、销售量等领域的有力工具，该方法对于本题的年度数据变化具有较好的解释性和预测能力。

### 3.3.2 模型计算

本文将1990-2014年的数据作为训练集，2015-2018年数据作为验证集，2019-2023年数据作为测试集。在ARIMA模型中，自回归和滑动平均参数是需要调优的超参数，对此本文先通过网格寻优方法，定义超参数的潜在参数空间，再通过K折交叉验证方法进行超参数寻优，得到参数范围内的最优值。

### 3.3.3 模型结果与评价

在训练好ARIMA模型后，本文在测试集上计算MSE值，评价模型预测值与实际值的平均误差，计算结果为403.28。这说明模型在很大程度上把握了数据分布规律，并且具有不错的预测能力；但是预测结果与现实存在差异，由于现实世界的复杂性和影响卫生医疗机构数量的因素过多，单一从历史数据角度看待卫生医疗机构数量变化会遗漏很多影响医疗卫生机构的不确定性因素。



### 3.3.4 影响因素分析

本文在广泛吸取他人研究成果的基础上，系统的总结了以下潜在的影响因素：在人口结构方面，死亡率、总人口数和总负担系数；在经济水平方面，人均生产总值和人均可支配收入。

接下来本文将尝试探讨上述因素对医疗卫生资源配置的潜在影响。在人口结构方面，死亡率的变化直接反映了一个地区健康状况的改善或恶化，如果死亡率较高，说明疾病负担较重，可能需要增加医疗资源以应对慢性病或急性疾病；如果死亡率下降，表明医疗卫生系统可能在改善，但仍需关注老年人口或特殊人群的健康需求。高死亡率区域可能需要更多的基础医疗设施和急救资源，而低死亡率区域可能更关注慢性病管理和预防措施。总人口数影响医疗资源的需求总量，人口多的地区需要更多的医院、诊所、医生等资源，以满足更大的人口基数的健康需求；人口增长较快的地区可能需要扩建医疗设施和增加医疗人员，而人口减少的地区可能面临资源闲置或需要优化资源分配。总负担系数是指一个地区在医疗支出、疾病负担等方面的总体压力，较高的负担系数表明医疗需求较高或医疗支出较重，可能需要更加优化的资源配置和政策支持；高负担系数的地区可能需要优先分配资源，用于改善医疗服务质量、增加疾病预防和健康教育等方面。

在经济水平方面，人均GDP反映了一个地区的经济发展水平，经济水平较高的地区通常能够支出更多的资金用于医疗卫生系统，从而提供更高质量的医疗服务；经济发展水平高的地区可能会优先投资高端医疗设备和技术，引进高级医疗人才，而经济水平较低的地区则需要更多基础设施和公共健康服务。人均可支配收入直接影响个人和家庭对医疗服务的支付能力，收入较高的家庭可以承担更多的自费医疗服务，而收入较低的家庭可能依赖公共医疗系统；人均可支配收入较高的地区可能更多地依赖私营医疗机构和自费服务，而收入较低的地区可能需要更多的公共医疗服务和补助。

## 3.4 问题四模型的建立与求解

本文综合考虑了当前医疗卫生资源配置、人口结构变化、经济发展水平政策导向等因素，充分借鉴了其它学者的研究成果，系统性的提出了我国医疗卫生机构测度体系，分为医疗卫生资源配置、人口结构、经济水平三个子层面，共含有八个指标，以此来进一步分析影响医疗卫生机构数量的关键因素，并且对中国未来10年（2024年至2033年）的医疗卫生机构数量进行宏观预测。

### 3.4.1 模型建立

在时间序列预测领域中，传统统计模型如自回归积分滑动平均（ARIMA）模型曾经占据主导地位。然而，随着深度学习技术的快速发展，长短期记忆网络（Long Short-Term Memory, LSTM）因其出色的表现和处理非线性数据的能力，成为研究和应用的热点。本文将采用LSTM模型进行进一步的分析和预测。

LSTM是一种特殊的递归神经网络（RNN），最初由Hochreiter和Schmidhuber在1997年提出，旨在解决普通RNN在长序列数据中存在的梯度消失和梯度爆炸问题。LSTM通过引入记忆单元和门控机制来有效地捕捉长时依赖关系。

LSTM由一系列重复的单元（cell）组成，其中每个单元包含三个门：输入门、遗忘门和输出门。这些门通过输入当前时刻的输入和先前时刻的隐藏状态来控制信息的流进和流出，从而维护和更新细胞状态。

相较于之前采用的ARIMA，LSTM模型具有以下优越性：

- 1) 非线性处理能力：LSTM无需假设数据是线性和平稳的，可以自动学习数据中的复杂非线性关系，而ARIMA设计之初针对的是线性时间序列。
  - 2) 长时依赖捕获：通过其独特的门控机制，LSTM能够有效记忆和利用长期的序列依赖信息，这是ARIMA所难以做到的，后者主要依赖于自回归结构处理短时依赖。
- 尽管ARIMA在处理小规模和线性数据时有其优势，但由于更多特征的增加，数据复杂度和维度变得更加复杂，这时LSTM提供了一种更加灵活和精确的解决方案。

### 3.4.2 模型计算

LSTM的基本计算方式如下：

- 遗忘门计算：确定细胞状态中哪些信息需要丢弃。通过激活函数  $\sigma$  应用于当前输入和前一时刻的隐藏状态，这一过程的计算公式为：

$$f = \sigma(W_{fx} + U_{fh} + b_f)$$

- 输入门计算：决定当前时刻输入的信息要存入多少到细胞状态。包括输入门向量和候选状态的计算：

$$i = \sigma(W_{ix} + U_{ih} + b_i)$$

- 更新细胞状态：通过前一时刻的细胞状态与当前候选细胞状态进行结合，更新新细胞状态：

$$c = f \circ g + i \circ \tanh(W_{cx} + U_{ch} + b_c)$$

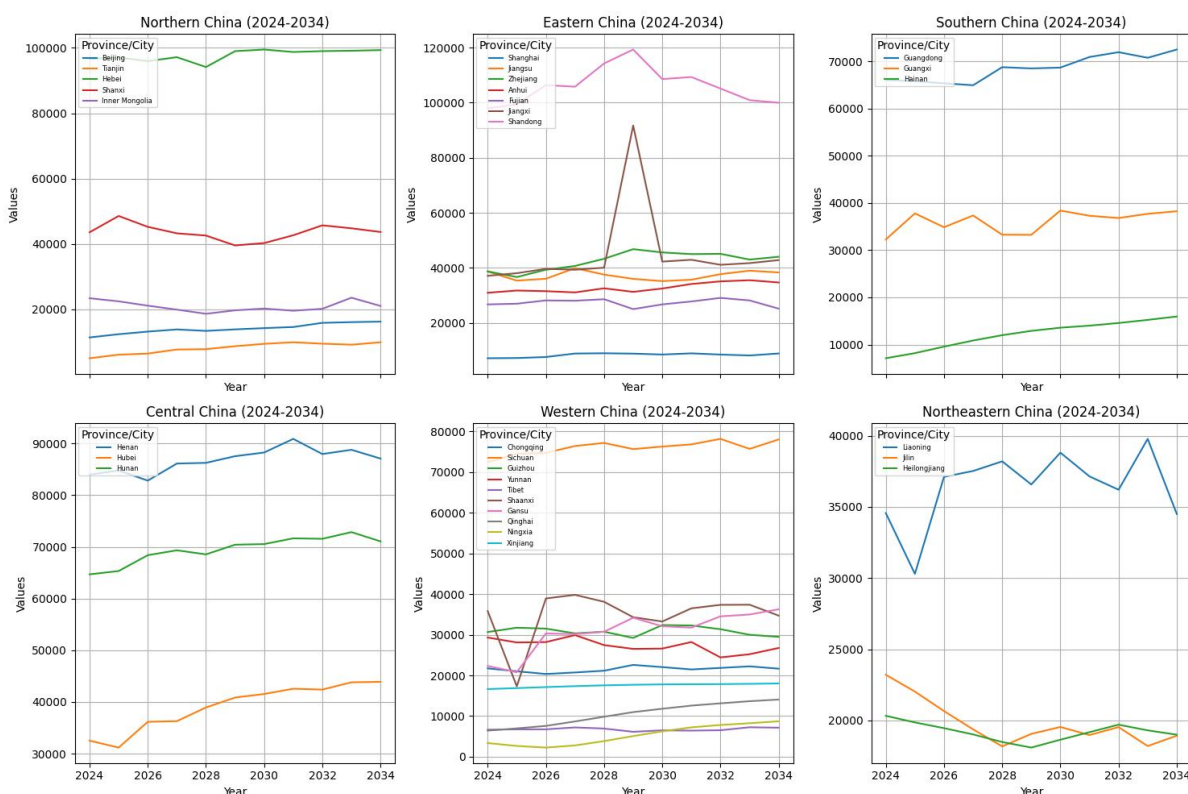
- 输出门计算：确定输出何种信息到隐藏状态：

$$o = \sigma(W_{ox} + U_{oh} + b_o)$$

本文将1990-2014年的数据作为训练集，2015-2018年数据作为验证集，2019-2023年数据作为测试集。在LSTM模型中，隐藏单元数、LSTM层数、学习率、批量大小、序列长度和正则化参数是需要调优的超参数，对此本文先通过网格寻优方法，定义超参数的潜在参数空间，再通过K折交叉验证方法进行超参数寻优，得到参数范围内的最优值。

### 3.4.3 模型结果与评价

在训练好LSTM模型后，本文在测试集上计算MSE值，评价模型预测值与实际值的平均误差，计算结果为235.69。这说明模型在考虑了更多潜在因素并结合LSTM算法相比于只考虑历史数据的ARIMA效果得到了很大提升，很好的把握了数据分布规律，并且预测能力具有可信力；预测结果虽然与现实存在较小差异，由于现实世界具有复杂性并且医疗卫生机构数量受很多情况共同影响，较小的差异可以被理解。



### 3.4.4 启示与挑战

从2024-2034年的预测数据来看，绝大多数省份医疗卫生机构数量保持不变或稍有增长，说明我国医疗卫生机构建设持续稳定健康发展，没有过多应急需求且保持一定的规模水平。空间分布上看，医疗机构数量与地方的人口数关联密切，广东、山东、河南、四川等人口大省的医疗卫生机构数量显著多于其它省份，西部地区普遍数量较少；同时经济较为发达地区的密度较高，比如北京、上海、天津、重庆等地，虽然人口较少，但也保持了一定规模的医疗卫生机构。总体来看，我国未来十年医疗卫生机构数量总体处于平稳发展态势，分布上显著受人口影响，经济因素也起到了一定作用。

本文认为我国医疗卫生机构建设还存在一些挑战，总结来说：

- 1) 在常住人口较少、经济欠发达地区的医疗卫生机构数目较少，医疗卫生机构数量是公共卫生发展的体现，为了贯彻协调、共享的新发展理念、改变区域发展不协调的局面，政府可以推进少人口、欠发达地区的公共卫生建设，从而保障人民的生命健康，提高人民的生活水平，在一定程度上这也能吸引更多从业者、缓解地区人口流失。
- 2) 少数省份在特定年份的医疗卫生机构数量预测结果显示出了突然性的波动，比如上海在2029年数量激增而后又下落、陕西在2025年数量突然下落再回升等。这可能意味着该地区突然受特定事件影响（比如医疗卫生政策变化、疫情等突发公共卫生事件）而导致剧烈变化，政府应对此加强关注，并做好预防措施。
- 3) 华中地区和华南地区总体的医疗卫生机构数量呈现增长态势，说明未来该地区的医疗卫生需求将会增加，政府应考虑该地区普遍增长的因素，积极推进地区医疗卫生建设，满足人民群众的切身需求，推动医疗卫生建设稳定向好发展。

## 四、模型反思与提升

### 4.1 模型的优缺点

本文综合使用了 PCA、莫兰指数、ARIMA 和 LSTM 模型，每个模型各有其优势和局限性，反思总结如下：

#### 4.1.1 PCA

##### 1. 优点

- 降维能力强：PCA 可以有效地将高维数据降低到较低的维度，同时保留尽可能多的原始数据的信息。这样不仅减少计算负担，还能够帮助数据可视化。
- 无监督学习：作为一种无监督学习方法，PCA 不需要标签数据，这使其在处理大量未标记数据时非常有用。

##### 2. 不足

线性假设：PCA 基于线性假设，将数据变换为新的正交基。然而，许多实际场景中的数据具有非线性特征结构，可能无法通过线性方法有效捕捉。

特征解释性：PCA 提取的主成分方向是数学上的最优方向，但这些方向不一定具有现实意义，导致特征解释难度增加。

#### 4.1.2 莫兰指数

##### 1. 优点

- 检测能力较好：莫兰指数是衡量空间数据集中的自相关性或空间依赖性的工具，能够较为准确的从数据集中揭示地理现象的聚集现象和空间结构。
- 简单直观：莫兰指数通过单一的指标提供对数据空间自相关性的量化描述，值范围在-1 到 1 之间，易于解释。

##### 2. 不足

- 边界效应：在有限的空间区域上，边界效应可能影响莫兰指数的准确性，尤其在小样本中。

#### 4.1.3 ARIMA 模型

##### 1. 优点：

- 理论基础扎实：基于统计学理论，有稳固的数学基础，对线性时间序列数据的建模表现出色。
- 可解释性强：由于依赖于参数化模型，ARIMA 的结果能够较容易地进行解释。

##### 3. 不足：

- 线性假设限制：由于其建模的线性特点，ARIMA 在面对复杂的非线性时间序列数据时性能欠佳。
- 数据需求高：模型通常需要对时间序列进行差分处理以达到平稳性，并且其参数选择和优化需要较长的历史数据。

#### 4.1.4 LSTM 模型

1. 优点：
  - 处理非线性关系：LSTM 能够捕捉到时间序列中复杂的非线性关系，适合用于处理多维数据。
  - 长期依赖捕捉：通过其记忆门控机制，LSTM 能够有效地捕捉和学习长时间的依赖关系，这在许多复杂序列预测任务中表现优越。
2. 不足：
  - 计算复杂性：LSTM 的训练过程比传统统计模型复杂，需要调参数的工作量通常更大，而且容易过拟合。
  - 可解释性差：由于模型的高度复杂性，其内部运作机制难以人理解，这对需要解释型分析结果的应用场景提出挑战。

#### 4.2 未来的改进方向

- 1) 对于传统 PCA 模型，可以使用核 PC 和自编码器等非线性降维方法，增强对数据的非线性结构的捕获能力，提高在医疗卫生机构数量预测这一复杂问题的分析能力。
- 2) 对莫兰指数引入边界效应修正技术，或增大数据集的范围，以此降低边界对结果的偏差。
- 3) 提升医疗卫生机构数量预测的效果可以通过结合 ARIMA 和 LSTM 模型的优势，构建混合模型来实现。这样的组合可以同时利用 ARIMA 的线性处理能力和 LSTM 的非线性建模能力，从而在不同时间尺度和不同特征的影响下，提供更为准确的预测结果。这种协同的方法不仅能够提高模型预测的精度，还可以增强结果的稳定性和鲁棒性，最终为医疗资源的合理规划和配置提供更可靠的数据支持。

### 五、讨论与政策建议

本文基于对 1990 年至 2023 年间中国各地区医疗卫生机构数量年均增长率和时间趋势的深入分析，以及对不同地区医疗资源配置差异的揭示，制定有效的政策建议对于优化未来医疗资源分配和提升公共卫生服务水平具有重要意义。本研究提出如下政策建议，以供决策者参考：

首先，要高度重视经济欠发达地区的医疗资源配置问题。分析结果显示，区域经济水平是医疗资源配置的主要决定因素之一，尤其是在欠发达地区，抚养比和人口健康状况对资源分配的影响尤为显著。为此，政府应增加对这些地区的财政投入，不仅要在基础设施建设上加大力度，如建设更多的医疗机构和增加卫生技术人员的配备，还需改善现有设施的质量和服务水平。这样做不仅能够缩小区域间的医疗资源差距，提升欠发达地区居民的健康水平，也有助于促进这些地区的社会经济发展，吸引和留住人才。



其次，针对医疗资源过于集中的现象，特别是在一些发达地区，政府应采取措施鼓励医疗资源的合理分布。通过政策支持，推动医疗资源向偏远和资源匮乏地区倾斜，例如提供财政激励、技术支持和人才引进政策，以促进资源在全国范围内的均衡分布。此外，通过提高资源配置的透明度和效率，优化现有资源的使用，确保公共与私人投资的最大化效益。

进一步，考虑到相邻省份之间存在显著的资源扩散效应，政府应鼓励跨区域合作，探索联动发展的新模式。这包括建立区域医疗协作网络，共享医疗技术和人力资源，强化区域疫病防控的联动响应，确保在公共卫生领域实现规模经济效应和知识的有效传播。地方政府可就医疗服务的互补与共享达成共识，推进区域内医疗服务的协调发展，缓解单一地区的资源过度集中或匮乏现象。

与此同时，对当前和未来可能影响医疗资源数量和分布的突发公共卫生事件和政策变化应密切关注。明确建立快速响应和调控机制，加强对医疗系统脆弱性的识别和防控能力，特别是在面对疫情等公共卫生紧急事件时，提高灵活应对能力。政府需从政策、技术、管理多方位入手，增强全国范围内对突发情况的抵御力，降低对医疗资源分配的经济和社会冲击。

最后，鼓励以数据驱动的决策支持，加强大数据和现代信息技术的应用，推动医疗资源配置的智能化和精准化转型。从科学预测和动态调整入手，构建基于数据分析的医疗资源优化分配平台，为政策制定提供可靠依据。

## 参考文献

- [1]陈小璐,刘婷,盘薇.我国基层卫生资源配置区域差异及空间分布趋势研究——以分级诊疗政策为切入点[J].卫生软科学,2024,38(08):13-17+22.
- [2]刘晖,王乐乐,寇丽圆,等.我国基层医疗卫生资源配置效率及影响因素研究[J].卫生软科学,2024,38(08):53-57.
- [3]庞瑞芝,刘磊,李倩楠.人口老龄化趋势下卫生资源配置结构性失衡与社会福利损失[J/OL].中国地质大学学报(社会科学版):1-15[2024-08-19].<https://doi.org/10.16493/j.cnki.42-1627/c.20240716.001>.
- [4]魏萱,魏颖,陈麒麟,等.2018—2021 年我国医疗卫生机构卫生人力资源配置公平性分析[J].中国初级卫生保健,2024,38(05):8-12.
- [5]邓雪,王灿友.组态视角下医疗卫生资源配置效率及提升路径研究[J].统计与管理,2024,39(04):83-90.DOI:10.16722/j.issn.1674-537x.2024.04.001.
- [6]郑浩渠,郭丹,刘雄新,等.珠三角城市医疗卫生资源配置公平性研究[J].基层医学论坛,2024,28(10):11-13+21.DOI:10.19435/j.1672-1721.2024.10.004.
- [7]刘晓彤.我国医疗卫生资源配置效率测度及区域差异分析[D].河南大学,2024.DOI:10.27114/d.cnki.ghnau.2023.000336.
- [8]叶林祥,左秀平.中国粮食产业高质量发展:水平测度、时空特征与路径选择[J].农村经济,2023,(11):64-73.

附录

1. 第一问 1990-12023 年各地区医疗机构的年平均增长率

地区 (省)	年均增长率 (%)
河南	7.16
河北	6.80
山东	6.49
江西	6.13
广东	6.07
山西	5.70
福建	5.65
甘肃	5.65
广西	5.59
吉林	5.44
青海	5.42
陕西	5.27
湖南	5.17
四川	4.99
内蒙古	4.98
重庆	4.92
贵州	4.61
云南	4.53
安徽	4.50
辽宁	4.45
宁夏	4.43
新疆	4.42
浙江	4.24
湖北	4.02
江苏	3.38
北京	2.85
黑龙江	2.68
海南	2.09
天津	1.78
上海	-0.60
西藏	5.70

2. 第三问 2024-2028 年各地区医疗机构的数量预测值

年份	2024	2025	2026	2027	2028
北京	13106.06472	12710.7257	13474.49317	13359.65392	13444.59071
天津	5473.438024	5105.669024	6023.407509	8684.641338	5766.509404
河北	93202.06719	92720.55377	93247.48969	85936.1081	92550.17709
山西	43139.11038	45699.12521	38813.23712	42173.2706	38780.36719
内蒙古	25628.74709	26284.77157	27090.94279	26235.31187	27440.24831
辽宁	33571.91122	34113.30022	34683.05503	34136.93555	35621.10083
吉林	24678.22206	25788.80693	27325.56352	27321.24194	27460.86883
黑龙江	21600.10358	21575.42718	24897.75287	24766.08735	24667.37951
上海	6287.431301	5599.435221	6248.159684	7848.874117	6231.005807
江苏	36420.63886	36645.9739	38657.89022	36755.31099	39226.89289
浙江	35641.81098	35758.89804	36751.49177	35434.84475	37500.61133
安徽	30889.27711	31324.49456	34493.52828	38628.19869	34536.49549
福建	30031.94127	30172.24171	35784.02316	30682.53294	30784.14232
江西	40147.52849	39712.69181	47884.56779	47938.98119	48546.60581
山东	90144.77936	90681.19353	91119.13096	81220.96388	91064.01821
河南	83491.40114	83470.28162	86144.32766	86345.33841	86415.3199
湖北	38322.35904	38297.4471	41725.49282	41476.62224	41658.79782
湖南	60649.34915	63853.90496	55688.0902	55369.25157	55443.98423
广东	61660.16361	61397.31134	67643.26008	67533.11414	67563.11195
广西	35128.76459	34924.66645	35729.81031	34359.1008	35344.91088
海南	5876.07267	6428.134707	7850.236069	15444.98884	7329.479581
重庆	23771.48258	23957.72262	25220.38421	25162.47285	25106.08537
四川	75598.7196	76010.25845	77458.09039	71403.11736	75447.84901
贵州	30427.37448	30432.8506	32319.51685	34749.62813	31373.55001
云南	29417.61062	28994.0875	29753.43985	27440.88688	29921.28054
西藏	6927.06508	6940.104804	6969.113654	6972.765191	6987.07203
山西	35077.56734	35448.91404	38543.59504	39403.30835	38300.18953
甘肃	21948.01468	21774.08771	26217.02511	23450.23657	26534.50645
青海	5599.652888	5147.724393	5513.597658	10434.98383	6465.872202
宁夏	4988.422357	4796.231126	5904.63496	9616.547292	5961.25655
新疆	16386.39597	15246.17857	19812.63751	19439.53771	19147.28582

3. 第四问 2024-2034 年各地区医疗机构的数量预测值

年份	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034
北京	11393	12383	13180	13843	13394	13859	14248	14581	15859	16079	16230
天津	5029	6109	6457	7703	7808	8711	9434	9915	9482	9174	9906
河北	96601	97056	95955	97153	94133	98999	99496	98741	99008	99126	99326
山西	43569	48569	45236	43256	42569	39546	40265	42659	45698	44789	43659
内蒙古	23393	22439	21117	19906	18633	19685	20236	19562	20156	23556	21025
辽宁	34568	30303	37113	37526	38202	36577	38814	37143	36208	39780	34505
吉林	23216	22033	20665	19379	18183	19062	19546	18980	19537	18209	18933
黑龙江	20332	19867	19464	19023	18490	18099	18654	19182	19708	19310	19005
上海	7202	7280	7647	8917	9012	8882	8546	8980	8537	8209	8933
江苏	38680	35391	36062	39907	37570	36020	35201	35714	37727	38987	38365
浙江	38737	36680	39331	40707	43316	46816	45621	45013	45109	43034	44047
安徽	30981	31788	31558	31100	32586	31307	32512	34164	35107	35532	34709
福建	26750	27004	28195	28084	28634	25042	26754	27821	29109	28194	25209
江西	37133	38076	39657	39337	40076	91716	42287	42952	41127	41738	42832
山东	97950	99409	106371	105789	114230	119293	108558	109304	105121	100918	99966
河南	83969	84781	82821	86124	86256	87546	88261	90894	87963	88784	87084
湖北	32521	31180	36167	36294	38941	40860	41552	42563	42400	43811	43894
湖南	64694	65333	68397	69345	68554	70427	70543	71656	71565	72861	71063
广东	65997	65779	65336	64923	68750	68486	68662	70905	71923	70737	72487
广西	32254	37782	34849	37335	33267	33245	38378	37273	36802	37688	38232
海南	7104	8185	9558	10857	11988	12897	13573	14007	14565	15228	15945
重庆	21727	21015	20351	20719	21152	22587	22043	21466	21839	22231	21650
四川	72429	74481	74747	76429	77189	75668	76296	76833	78186	75722	78040
贵州	30672	31736	31506	30334	30752	29220	32388	32280	31361	29996	29485
云南	29315	28105	28195	29915	27453	26521	26604	28214	24425	25206	26750
西藏	6649	6729	6737	7205	6923	6128	6487	6410	6508	7246	7143
山西	35824	17312	38932	39817	38109	34315	33259	36500	37357	37389	34707
甘肃	22334	20788	30321	30159	30737	34187	32085	31761	34525	34979	36260
青海	6408	6971	7578	8674	9833	10970	11796	12569	13132	13653	14063
宁夏	3358	2646	2241	2774	3836	5072	6237	7222	7806	8241	8707
新疆	16633	16875	17127	17360	17557	17695	17782	17822	17854	17924	18023