基于聚类分析和主成分分析的城市空气质量评价

褚心雨

北京林业大学经济管理学院　北京

【摘要】随着城市化的发展，人们生活水平也日益提升，但我国大气污染形势严峻，空气质量并不乐观。空气质量问题成为我国面临的重点问题之一，影响着人们的生活与健康。针对全国重点城市空气质量问题，本文采用 K 均值聚类分析、主成分分析的方法对数据进行实验分析，得出了影响空气质量的主要因素。研究表明：根据聚类分析得到的类中心表，部分类别中可吸入颗粒物（PM_{10}）浓度显著偏高，污染较为严重。同时，主成分分析的第一主成分载荷显示，PM_{10} 和细颗粒物（$PM_{2.5}$）是影响空气质量的关键指标。进一步对比 K 均值聚类的分类结果与主成分分析的综合得分：空气质量被划分为优的城市在主成分综合得分中普遍排名靠前；而被归为差类别的城市，其综合得分则普遍靠后。这种趋势表明，聚类分析与主成分分析的结果在城市空气质量排序上具有一致性，这说明本文构建的空气质量评价模型的合理性和有效性。

【关键词】聚类分析；主成分分析；空气污染物；空气质量

【基金项目】华南植物资源研发中心技术推广平台科技创新专项-城市森林生物多样性提升技术研究与示范应用项目

【收稿日期】2026 年 8 月 20 日 【出刊日期】2026 年 11 月 18 日 【DOI】10.12208/j.aee.20260007

Urban air quality assessment based on cluster analysis and principal component analysis

Xinyu Chu

School of Economics and Management, Beijing Forestry University, Beijing

【**Abstract**】With the development of urbanization, people's living standards have been steadily improving. However, the air pollution situation in our country remains severe, and air quality is not optimistic. Air quality has become one of the key issues facing our country, affecting people's daily lives and health. To address the air quality problems in key cities nationwide, this study conducts experimental analysis on the data using K-means clustering and principal component analysis (PCA) methods to identify the main factors affecting air quality. The research shows that: according to the cluster centroid table obtained from the clustering analysis, some categories have significantly higher concentrations of inhalable particulate matter (PM_{10}), indicating severe pollution. Meanwhile, the first principal component loadings from PCA indicate that PM_{10} and fine particulate matter ($PM_{2.5}$) are key indicators affecting air quality. Further comparison of the K-means clustering results with the comprehensive scores from PCA shows that cities classified as having good air quality generally rank high in the PCA comprehensive scores, while cities classified as having poor air quality rank low. This trend indicates that the results from the clustering analysis and PCA are consistent in the ranking of urban air quality, demonstrating the rationality and effectiveness of the air quality evaluation model constructed in this study.

【**Keywords**】Cluster analysis; Principal component analysis; Air pollutants; Air quality

1 引言

当前，我国大气污染形势严峻。以可吸入颗粒物、细颗粒物为特征污染物的区域性大气环境问题日益突出，损害人民群众的身体健康，这最终将影

作者简介：褚心雨，女，硕士研究生，主要研究方向：资源环境统计

响社会的和谐稳定。随着我国工业化、城镇化的深入推进，能源和资源消耗持续增加，大气污染防治压力继续加大。尤其是近几年我国出现大面积的大气污染事件，直接反应了大气污染的严峻形势。因此，详细了解我国大气质量状况，发现污染原因，深入研究大气污染对人们健康的影响是十分必要和紧迫的任务[1]。影响空气质量的因素多而复杂，只有找到最主要的因素，才可以较好地解决空气质量问题。

本文借助 SPSS 软件，使用来源于 2020 年《中国统计年鉴》[2]的数据进行统计分析。首先用聚类分析对样品进行量化分类，K 均值法能够快速聚类，以距离的远近亲疏为标准；之后用主成分分析提取最主要的主成分从而分析最主要的影响因素，通过“降维”来进行数据的综合评价及排序，在“降维”过程中原数据的大部分信息能被保留，不失数据的真实性[3]。

2 数据来源与研究方法

2.1 数据来源

数据来源于 2020 年《中国统计年鉴》城市空气质量情况，31 个主要城市，6 个空气影响因素（X1、X2、X3、X4、X5、X6）分别是：二氧化硫（SO_2）年平均浓度（μg/m³）、二氧化氮（NO_2）年平均浓度（μg/m³）、可吸入颗粒物（PM_{10}）年平均浓度（μg/m³）、一氧化碳（CO）日均值第 95 百分位浓度（μg/m³）、臭氧（O_3）日最大 8 小时第 90 百分位浓度（μg/m³）、细颗粒物（$PM_{2.5}$）年平均浓度（μg/m³），这 6 个因素也是我国环境空气质量标准全国实施的基本项目。

2.2 研究方法

2.2.1 K 均值聚类分析

Macqueen 于 1967 年提出了 K-均值法，这种聚类方法的思想是把每个样品聚集到其最近形心（均值）类中[4]。K 均值聚类的基本原理，是将数据中每个城市当成一个样品，每个城市的空气质量指标则为样品的要素，在它的最简单说明中，这个过程由下列三步组成[5]：

（1）把样品粗略分成 K 个初始类。

（2）进行修改，逐个分派样品到其最近均值类中（通常用标准化数据或非标准化数据计算欧氏距离）。重新计算接受新样品的类和失去样品的类的形心（均值）。

（3）重复第（2）步，直到各类无元素进出。

2.2.2 主成分分析

主成分分析是从多个原始变量中找出较少的主成分且能反映原来信息的真实性。原始变量间存在或强或弱的相关性，每个原始变量所反映的信息也不一样，这些信息不同程度的决定着它们在整体中的重要程度。主成分分析法通过线性组合的方式来达到“降维”，能快速找到有影响力的主要因素。

（1）数据标准化，建立相关系数矩阵 R，由于变量间的单位可能不同，避免在主成分分析上带来影响，将原始变量数据标准化：

$$X_i^* = \frac{X_i - E(X_i)}{\sqrt{D(X_i)}},\ i = 1,\ 2,\ \cdots,\ p$$

P：p 个变量数；X_i：第 i 个变量。

（2）计算 R 的特征根 λ_i 以及相应的特征向量 T_i。

（3）主成分的个数由累计方差贡献率确定，方差贡献率：

$$\varphi_i = \frac{\lambda_i}{\sum_{i=1}^{p} \lambda_i}$$

累计方差贡献率：

$$\varphi_k = \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{p} \lambda_i} k < p;$$

唯有累计方差贡献率达到比较高的百分数，才可确定有 k 个主成分，进而确定主成分表达式及主要影响因素。一般情况下，主成分的个数由特征值大于 1 的个数确定，本文采用此方法进行计算。

（4）主成分表达式及加权综合主成分表达式，解释主成分的含义并对实际问题的各个样品，按照每个主成分进行排序，综合评价。

主成分表达式：

$$Y_i = (T_i)X^*;\ i = 1,\ 2,\ \cdots,\ k$$

加权综合主成分表达式：

$$Y = \sum_{i=1}^{k} \frac{\lambda_i}{\sum_{m=1}^{p} \lambda_m} Y_i$$

2.3 指标选取

本文选取的指标的原因介绍如下：第一，悬浮颗粒物污染对人体健康的影响。主要与化学成分和其粒径大小有关。而颗粒物的大小决定了它们最终在呼吸道中的位置，较大的颗粒物会被纤毛和黏液过滤无法通过鼻子和咽喉，PM_{10} 可以穿过这些屏障到达支气管和肺泡，$PM_{2.5}$ 则更具有穿透力，可能抵达支气管壁，并干扰肺内的气体交换。第二，二氧化硫污染与

人体健康。在大气中，二氧化硫通过复杂的化学变化进而氧化形成硫酸雾或硫酸盐气溶胶，是环境酸化的重要前驱物。第三，氮氧化物污染与人体健康。在氮氧化物中，二氧化氮毒性最大，与一氧化氮相比，其毒性是它的 4-5 倍。大气中氮氧化物主要来自煤和石油燃烧的废气以及汽车废气。第四，臭氧污染与人体健康。臭氧污染是一种典型的大气二次污染，具有强氧化能力的臭氧具有腐蚀性，其对人体健康的危害主要体现在：强烈刺激呼吸道、造成神经中毒、致使人的皮肤起皱出现黑斑、破坏人体免疫机能、加速衰老、致使胎儿畸形。第五，一氧化碳进入人体后能和血红蛋白结合形成碳氧血红蛋白。当血红蛋白与一氧化碳结合后会导致血红蛋白与氧的结合能力差，导致病人出现低氧血症[6]，本文使用的数据如表 1 所示。

表 1　2019 年我国主要城市空气污染物浓度表（单位：mg/m^3）

城市	SO_2	NO_2	PM_{10}	CO	O_3	$PM_{2.5}$
北京	4	37	68	1.4	191	42
天津	11	42	76	1.8	200	51
石家庄	16	46	118	2.4	206	63
太原	22	50	107	1.9	186	56
呼和浩特	15	39	71	2.2	146	37
沈阳	21	36	77	1.9	155	43
长春	11	34	64	1.3	134	38
哈尔滨	17	32	67	1.4	116	42
上海	7	42	45	1.1	151	35
南京	10	42	69	1.3	181	40
杭州	7	41	66	1.1	181	38
合肥	6	42	68	1.2	168	44
福州	5	22	42	1.4	138	24
南昌	9	34	68	1.4	151	35
济南	15	45	106	1.6	201	55
郑州	9	45	99	1.6	194	58
武汉	9	44	71	1.5	183	45
长沙	7	33	57	1.3	171	47
广州	7	45	53	1.2	178	30
南宁	9	29	53	1.3	126	30
海口	5	13	32	0.9	144	17
重庆	7	40	60	1.2	157	38
成都	6	42	68	1.1	160	43
贵阳	10	21	47	0.9	125	27
昆明	11	31	45	1.0	134	26
拉萨	6	19	31	0.9	129	12
西安	9	48	96	1.7	1 66	57
兰州	18	50	79	2.5	151	36
西宁	17	37	59	2.3	129	34
银川	15	37	68	2.0	147	31
乌鲁木齐	8	42	84	2.5	127	50

3 结果与分析

3.1 K 均值聚类分析

本文使用 K 均值法对主要城市空气质量数据进行聚类分析。

3.1.1 数据标准化

实验数据的影响因素在单位上虽大部分是一致

的，但所测量的时间段不同，有年平均、日均值、最大 8 小时等，且各因素所要求的标准含量不一致，存在着偏差，在聚类分析后样品离类中心的距离与各因素类中心无法直接对比以及分析哪类较好。为防止这些偏差对实验造成影响，需要使用 SPSS 对数据进行标准化处理。

3.1.2 实验说明

运用 K 均值法对重点城市空气质量数据进行聚类，通过实验分为 3 类：优、良、差。通过 SPSS 对标准化的数据进行聚类分析。

以下是符号说明：X1 代表二氧化硫（SO_2）对比值，X2 代表二氧化氮（NO_2）对比值，X3 代表可吸入颗粒物（PM_{10}）对比值，X4 代表一氧化碳（CO）对比值，X5 代表臭氧（O_3）对比值，X6 代表细颗粒物（$PM_{2.5}$）对比值。

3.2 实验结果

各城市所属类别以及各城市距类中心的距离、最终类中心表、每个聚类中的案例数见表 2。

表 2 各城市所属类成员及距类中心距离表

城市	类别	距离	城市	类别	距离
北京	2	1.940	武汉	2	1.283
天津	3	1.540	长沙	2	1.503
石家庄	3	1.542	广州	2	1.786
太原	3	1.672	南宁	1	1.244
呼和浩特	2	1.665	海口	1	1.457
沈阳	2	2.344	重庆	2	1.119
长春	2	1.206	成都	2	1.429
哈尔滨	2	2.225	贵阳	1	.809
上海	2	1.674	昆明	1	1.207
南京	2	1.135	拉萨	1	1.206
杭州	2	1.551	西安	.	.
合肥	2	1.379	兰州	2	2.802
福州	1	.914	西宁	2	2.348
南昌	2	.880	银川	2	1.519
济南	3	.627	乌鲁木齐	2	2.614
郑州	3	1.285			

表 3 最终类中心表

	类别		
	1	2	3
X1	-.60251	-.00694	.81536
X2	-1.63579	.21947	.89694
X3	-1.25859	-.08410	1.56601
X4	-.95603	.09969	.69587
X5	-.98713	-.07527	1.47057
X6	-1.40637	-.00965	1.43137

表 4 每个聚类中的案例数表

类别	个数
1	6
2	19
3	5
有效个数	30
错失个数	1

由表 3 可知最终各因素的类中心，其中主要是类别 2 因素 2 较高，类别 3 因素 3、5、6 超标。由这 31 个重点城市可看出我国绝大部分的城市空气污染物 NO_2（二氧化氮）以及颗粒物（PM_{10}）超标且部分超标严重。由表 4 可知，城市所属类别 1 有 6 个城市，类别 2 有 19 个城市，类别 3 有 5 个城市。

各因素最终类中心分析可直接通过各类别超标的个数判断各类别，其中类别 1 为优、类别 2 为良、类别 3 为差。

由于已选择聚类以使不同聚类中个案之间的差异最大化，因此 F 检验只应该用于描述目的。实测显著性水平并未因此进行修正，无法解释为针对“聚类平均值相等”这一假设的检验。由表 5 可知，这六个变量的 sig 值均小于 0.05，通过显著性检验，即这六个变量对分类的贡献显著。

3.3 主成分分析

本文用主成分分析对重点城市空气质量数据进行综合评价和排序以及得出主成分，并计算各主成分因子得分，综合得分。具体方式为，直接用原始数据的相关系数矩阵 R 代替标准化后的数据的协方差阵。

3.3.1 相关阵计算数据标准化

城市空气质量各指标数据差别大且存在不同的物理量纲，可以把原始数据的相关系数矩阵当作标准化的数据实验。

3.3.2 相关阵计算主成分实验结果

将表 1 中的数据录入 SPSS 表，选择从 SPSS 默认设置的相关阵出发对数据进行主成分分析。

由表 6 可以看出，除了 SO_2 与 O_3 的相关性不强之外，其他变量之间均存在较强的相关关系，因此原始数据适合做主成分分析。

相关系数矩阵的特征根与方差贡献率见表 7。

表 5 方差分析表

	聚类		误差		F	显著性
	均方	自由度	均方	自由度		
Zscore（SO_2）	2.750	2	.903	27	3.044	.064
Zscore（NO_2）	10.474	2	.284	27	36.910	.000
Zscore（PM_{10}）	10.921	2	.236	27	46.373	.000
Zscore（CO）	4.045	2	.806	27	5.015	.014
Zscore（O_3）	8.383	2	.453	27	18.503	.000
Zscore（$PM_{2.5}$）	11.021	2	.213	27	51.828	.000

表 6 相关性矩阵表

		SO_2	NO_2	PM_{10}	CO	O_3	$PM_{2.5}$	SO_2
相关性	SO_2	1.000	.366	.558	.649	.014	.362	
	NO_2	.366	1.000	.752	.524	.617	.757	
	PM_{10}	.558	.752	1.000	.644	.627	.908	
	CO	.649	.524	.644	1.000	.128	.497	
	O_3	.014	.617	.627	.128	1.000	.672	
	$PM_{2.5}$	.362	.757	.908	.497	.672	1.000	

表 7 总方差解释表

	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 %
1	3.791	63.181	63.181	3.791	63.181	63.181
2	1.281	22.344	85.526	1.281	22.344	85.526
3	.360	4.998	90.523			
4	.279	4.653	95.176			
5	.233	3.882	99.058			
6	.057	.942	100.000			

由表 7 可知，本案例保留了两个主成分，它们解释了全部变量总方差的 85.526%，说明这两个主成分能有效代表原来的六个指标评价主要城市的空气质量。

表 8 结果使用主成分分析法提取 2 个成分获得。对成分矩阵表中成分 1 和成分 2 所对应的两列元素分别除以第一个和第二个特征根的平方根 √3.791，√1.281，得到两个主成分的变换系数，如表 9 所示。

由表可得两个主成分的线性表达式如下：

$$Y_1 = 0.30405X_1^* + 0.443235X_2^* + 0.493054X_3^* + 0.366195X_4^* + 0.339488X_5^* + 0.467374X_6^*$$

$$Y_2 = 0.679534X_1^* - 0.16692X_2^* - 0.0158X_3^* + 0.539281X_4^* - 0.64793X_5^* - 0.21927X_6^*$$

其中，X^*表示对原始变量标准化后的变量。

主成分的实际意义由各线性组合中系数较大的几个指标的综合意义来确定。第一主成分方差的贡献率 63.181%，占原信息的大部分，起最大作用。观察第一主成分的特征向量，主成分 Y_1 中各变量系数大小相当，PM_{10} 和 $PM_{2.5}$ 的系数最大，所以主成分 Y_1 主要考虑了空气中颗粒物的浓度含量水平。

第二主成分反映了原信息的 22.344%，影响力没有第一主成分高，主成分 Y_2 中 SO_2 和 CO 的系数最大，所以 Y_2 主要考虑了有毒性氧化物的浓度含量水平。这两个主成分从颗粒物浓度和有毒氧化物浓度两个水平来刻画空气质量状况，具有 85.526%的可靠性。

之后将标准化后的数据代入两个主成分的线性表达式，计算各样品的两个主成分得分，结果如表 10 所示。

将各样品的主成分得分在平面直角坐标系上使用 SPSS 软件画散点图并添加辅助线描出来，结果如图 1 所示。

由图 1 可知，分布在第一象限的地区是西宁、兰州、沈阳、银川、呼和浩特、乌鲁木齐、太原，说明这些城市综合来说空气中颗粒物浓度高，SO_2 等有毒氧化物浓度高，整体空气质量较差；分布在第三象限的是南昌、海口、重庆、上海、长沙、广州、成都、杭州；说明这些城市空气中颗粒污染物浓度低，SO_2 等有毒氧化物浓度低，整体空气质量较好。

3.3.3 从协方差阵出发计算主成分

表 11 给出了该次分析所保留的 m 个主成分从每个原始变量中提取的信息。主成分 NO_2 的贡献率为 82.64/54.232=0.661，可以看出主成分包含变量 O_3 的 99.9%信息，但是包含变量 SO_2 的信息最少。

表 8　主成分矩阵表

	成分	
	1	2
SO_2	.592	.688
NO_2	.863	-.169
PM_{10}	.960	-.016
CO	.713	.546
O_3	.661	-.656
$PM_{2.5}$	.910	-.222

表 9　变换系数矩阵表

	主成分	
	1	2
SO_2	0.30405	0.679534
NO_2	0.443235	-0.16692
PM_{10}	0.493054	-0.0158
CO	0.366195	0.539281
O_3	0.339488	-0.64793
$PM_{2.5}$	0.467374	-0.21927

表 10　主成分得分表

	第一主成分得分	第二主成分得分		第一主成分得分	第二主成分得分
北京	-0.016920084	-1.893935738	武汉	0.794725163	-1.075392333
天津	1.621224497	-0.955954379	长沙	-0.420714252	-1.107569114
石家庄	4.112880347	-0.059716905	广州	-0.581582962	-1.296953154
太原	3.510385623	0.768011539	南宁	-1.828773264	0.665272041
呼和浩特	0.668974459	1.692814282	海口	-3.927399009	-0.235599268
沈阳	1.158378393	1.908723686	重庆	-0.61881163	-0.84079267
长春	-0.788312107	0.499952133	成都	-0.238810788	-1.300133872
哈尔滨	-0.441669721	1.849830041	贵阳	-2.730874017	0.585591391
上海	-1.14335867	-0.775835915	昆明	-2.076339609	0.452267696
南京	0.339213495	-0.982025171	拉萨	-3.985795302	0.254988358
杭州	-0.196764051	-1.566280022	西安	1.88182772	-0.564211458
合肥	-0.020358825	-1.402972718	兰州	1.831339869	2.134607118
福州	-2.678583873	0.17288397	西宁	0.155300798	2.601841547
南昌	-0.64098516	-0.031863388	银川	0.127483616	1.592507817
济南	2.734254279	-0.800009226	乌鲁木齐	1.175595059	1.221296662
郑州	2.224464293	-1.511366529			

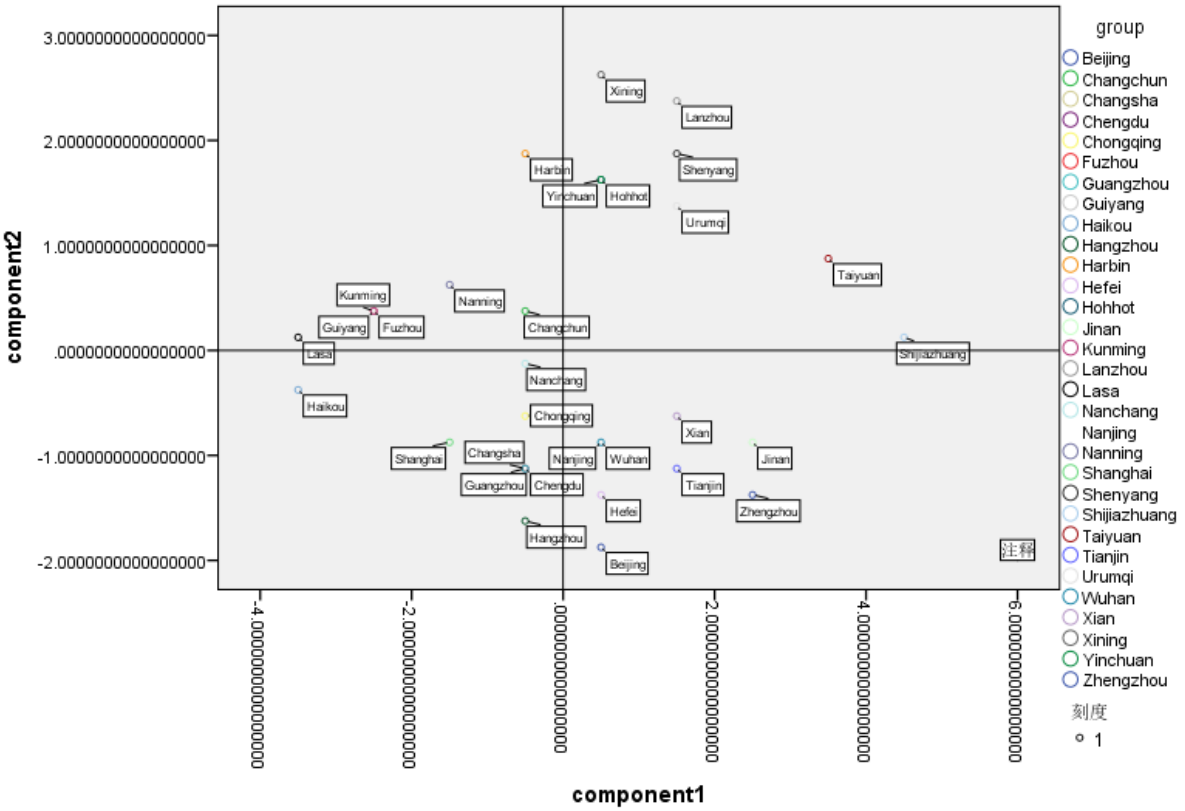


图 1　样品主成分得分散点图

表 11 公因子方差表

	原始		重新标度	
	初始	提取	初始	提取
SO_2	24.644	12.617	1.000	.512
NO_2	82.064	54.232	1.000	.661
PM_{10}	431.995	427.165	1.000	.989
CO	.238	.130	1.000	.547
O_3	693.747	693.316	1.000	.999
$PM_{2.5}$	136.990	121.893	1.000	.890

在分析协方差矩阵时，原始解与重新标度的解的初始特征值相同。表 12 给出了主成分解释原始变量、第二主成分总方差的情况。SPSS 默认保留大于特征根均值的特征根，本案例保留了两个特征根，分别是 1078.315 和 231.038，二者分别是第一主成分和第二主成分的方差。原始部分显示第一主成分和第二主成分解释了全部变量总方差的 95.596%，说明这两个主成分代表原来的六个指标评价主要城市的空气质量已经足够。重新标度部分可以看出第一、二主成分占所有主成分对原始变量方差贡献率总和的比值为 0.766。

由图 2 也可以看出，折线在因子个数为 2 的位置由陡峭突然变得平稳，因此选择两个主成分是合适的。

表 12 总方差解释表

		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（%）
原始	1	1078.315	78.728	78.728	1078.315	78.728	78.728
	2	231.038	16.868	95.596	231.038	16.868	95.596
	3	30.454	2.223	97.819			
	4	21.366	1.560	99.379			
	5	8.412	.614	99.993			
	6	.093	.007	100.000			
重新标度	1	1078.315	78.728	78.728	3.209	53.489	53.489
	2	231.038	16.868	95.596	1.389	23.143	76.632
	3	30.454	2.223	97.819			
	4	21.366	1.560	99.379			
	5	8.412	.614	99.993			
	6	.093	.007	100.000			

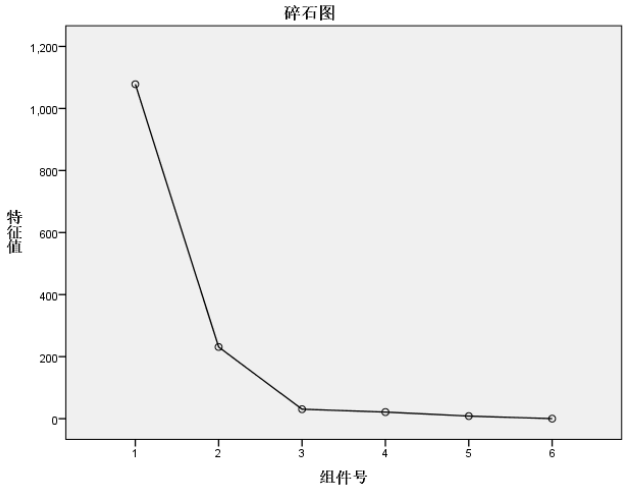


图 2 碎石图

表 13 成分矩阵表

	原始		重新标度	
	成分		成分	
	1	2	1	2
SO_2	1.348	3.286	.271	.662
NO_2	7.024	2.213	.775	.244
PM_{10}	18.110	9.960	.871	.479
CO	.191	.306	.391	.628
O_3	24.377	-9.953	.926	-.378
$PM_{2.5}$	10.242	4.122	.875	.352

表 13 结果由主成分分析法提取 2 个成分获得。将原始成分对应的列向量除以√λ，得到特征向量，进而得到主成分的表达式为：

$$Y_1 = 0.041050369X_1 + 0.213900441X_2 + 0.55150014X_3 + 0.005816484X_4 + 0.742347814X_5 + 0.311897539X_6$$

$$Y_2 = 0.216185146X_1 + 0.145592735X_2 + 0.655265994X_3 + 0.020131666X_4 - 0.654805466X_5 + 0.271185384X_6$$

根据成分矩阵主成分对应的因子载荷量可以看到 X_2、X_3、X_5、X_6 对应在第一主成分中占据较大比重，说明第一主成分主要综合了 NO_2、PM_{10}、O_3、$PM_{2.5}$ 在空气中的浓度水平。而 X_3、X_5 对应在第二主成分中占据较大比重，说明第二主成分主要综合了 PM_{10}、O_3 在空气中的浓度水平。

综合分析从相关阵出发、从协方差阵出发计算主成分，虽然第二种方法主成分对方差的解释率更高，但是不具有实际意义。因此考虑使用 3.3.2 中的结果进行综合分析。

4 讨论

本文是对全国重点城市来分析的，对个别城市的重点区域进行分析也是可以的，可将区域划分成多个样品进行监测并实验，使得实验更精确有效。本文对全国主要城市的空气质量评价是以一年的数据分析的，若要进行多年间的对比或预测，还需建立对比模型或预测模型，需考虑更多指标以及更优化的算法。

5 结论

从聚类结果的类中心表，可以看出类别中可吸入颗粒物（PM_{10}）超标严重。同样，从主成分分析的第一主成分可以看出可吸入颗粒物（PM_{10}）、细颗粒物（$PM_{2.5}$）是影响空气质量的主要因素。

通过 K 均值聚类各类别的数目来对应主成分分析结果对比：优等城市在主成分结果综合得分中都是靠前的，差等的城市在主成分结果综合得分都是靠后的。由此可见：聚类分析与主成分分析的实验结果在整体上是一致的，说明了所建立的对城市空气质量的评价模型是可行的。

城市空气质量评价是大气环境保护的一项重要工作，在实践中也起到重要的作用，只有对空气质量环境做出客观、正确的分析，找到影响空气质量的关键因素，才能有效防治空气污染。当前我国大气污染形势严峻，区域性大气污染问题突出，污染总体上北方重于南方，空气污染程度主要受空气中颗粒物浓度和有毒氧化物的影响。大气污染给人体健康带来极大的伤害，容易提高了死亡率，增加呼吸系统疾病的发生，对人体免疫系统造成极大的危害，大气环境问题迫在眉睫，各城市尤其是空气质量较差的城市应及时采取各项措施减少各空气污染物的浓度。

参考文献

[1] 姜少睿,薛志钢,李薇,等.我国环境空气质量状况及大气污染对健康的影响[J].华北电力技术, 2015(08):7-13.

[2] 2020 中国统计年鉴[J].统计理论与实践,2021(01):2.

[3] 魏臻,林芳.基于聚类分析和主成分分析城市空气质量评价[J].淮阴工学院学报,2018,27(03):86-96.

[4] MacQueen, J.C. Some Methods for Classification and Analysis of Multivariate Observations, the 5th Berkeley

Symposium on Mathematics, Statistics and Probability, 1967, 1: 281-297

[5] 何晓群.多元统计分析[M].北京:中国人民大学出版社, 2019.

[6] 鲁然英.城市环境空气质量及其评价方法研究[D].兰州大学,2006.

版权声明：©2026 作者与开放获取期刊研究中心(OAJRC)所有。本文章按照知识共享署名许可条款发表。
https://creativecommons.org/licenses/by/4.0/

OPEN ACCESS