

登记申请书

数据申请号	粤2025123103653		申请日	2026.02.24
数据名称	京津冀PM2.5空气质量多维度特征数据集(2015-2026)			
申请人	姓名或名称	钟正洋	申请人类型	<input checked="" type="checkbox"/> 自然人 <input type="checkbox"/> 法人或非法人组织
	联系人	钟正洋	联系电话	18072740843
	居民身份证件 号码或统一社 会信用代码	4412*****558		
	经常居住地或 营业所在地信 息	广东-广州市-花都区 -1		
数据简介	本数据集涵盖2015-2026年京津冀地区13个城市的PM2.5浓度数据，通过创造性特征工程处理，从原始3个基础字段扩展为32个多维度特征，包含时间模式、污染事件、空间关系、政策影响和健康风险等五个维度的衍生知识，为环境监测、政策评估和健康风险分析提供高价值数据支持。			
关键词	PM2.5, 京津冀, 空气质量, 特征工程, 污染事件, 政策评估, 健康风险, 时间序列, 空间分析		数据所属行业 分类	N77 生态保护和环境治理业
数据来源	<input type="checkbox"/> 自行生产 <input checked="" type="checkbox"/> 公开收集 <input type="checkbox"/> 协议取得 <input type="checkbox"/> 其他合法来源			
数据处理规则 简要说明	一、加工工具 - 主要工具: Python 3.12、pandas 2.2.0、numpy 1.26.0 - 辅助工具: Jupyter Notebook (数据探索)、VS Code (代码开发) 二、核心算法 1. 数据预处理: 线性插值 (缺失值处理)、异常值检测 (基于GB 3095-2012标准范围) 2. 时间特征: 7天/30天滑动平均 (移动窗口算法)、30天趋势分析 (线性回归)、同比变化计算 (百分比差异) 3. 污染事件识别: 基于规则的事件检测 (连续3天>75 μg/m³且呈上升趋势) 4. 空间特征: 区域统计 (均值计算)、偏离程度 (百分比差异)、排名算法 (降序排列) 5. 健康风险: AQI等级分类 (基于GB 3095-2012标准)、超标标志 (阈值判断) 三、适用标准 - GB 3095-2012《环境空气质量标准》(PM2.5浓度限值、AQI等级划分) - HJ 633-2012《环境空气质量指数(AQI)技术规定(试行)》(AQI计算方法) - 《数据安全法》《个人信息保护法》(数据处理合规性) 四、个人信息处理 本数据集不涉及个人信息, 仅包含城市级别的环境监测数			



数据当前使用状态	<input checked="" type="checkbox"/> 自用 <input type="checkbox"/> 已许可 <input type="checkbox"/> 已质押 <input type="checkbox"/> 其他 _____		
应用场景描述	<p>一、适用条件 - 数据完整覆盖2015-2026年京津冀地区 - 包含32个多维度特征，支持深度分析 - 符合GB 3095-2012等国家标准 二、适用范围 - 地域范围：京津冀地区13个城市（北京市、天津市、石家庄市等） - 时间范围：2015-2026年年度数据 - 业务范围：环境监测、政策评估、健康风险分析、科学研究、智慧城市建设 三、适用对象 - 环保部门：环境监测与预警 - 政策制定者：政策效果评估与优化 - 公共卫生部门：健康风险评估与防控 - 科研机构：环境科学研究 - 智慧城市建设者：环境数据智能化管理 四、解决的主要问题 1. 污染预警：识别污染趋势，提前预警污染事件，支持应急响应 2. 政策评估：量化政策实施效果，为政策调整提供数据支撑 3. 健康风险防控：评估公众暴露风险，制定针对性防护措施 4. 区域联防联控：分析区域污染传输，支持协同治理 5. 科研支持：为环境科学研究提供多维度数据，推动科学发展 6. 智慧城市管理：实现环境数据智能化管理，提升城市管理水平 五、应用价值 本数据集通过多维度特征工程，将原始数据转化为高价值知识，有效支撑</p>		
数据结构样例	<p>history_processed表字段说明清单 字段名：id 含义：数据记录的唯一标识主键 用途：区分每一条独立的监测数据记录，方便数据查询、去重及关联追溯（例如存证时对应单条数据的标识）。 字段名：城市 含义：PM2.5 监测数据对应的京津冀区域城市 / 地区名称 用途：作为数据的区域维度分类依据，支持后续按城市做污染情况对比分析。 字段名：日期 含义：PM2.5 数据的监测日期 用途：作为数据的时间维度分类依据，支持按日 / 月 / 年统计污染趋势，同时匹配存证中“数据时间范围”的信息项。 字段名：AQI 含义：空气质量指数（Air Quality Index），反映当日空气质量的综合指标 用途：辅助 PM2.5 数据做空气质量综合评估，补充环境分析的维度。 字段名：质量等级 含义：基于 AQI 划分的空气质量等级（如“优 / 良 / 轻度污染”等） 用途：直观体现当日空气质量水平，便于快速分类统计不同等级的天数占比。 字段名：PM2.5 含义：空气中 PM2.5 污染物的浓度值（单位通常为 $\mu\text{g}/\text{m}^3$） 用途：本数据集的核心监测指标，是环境分析、存证登记的核心数据内容。</p>		
数据格式	CSV数据库文件	数据量	52,481
数据更新频率	<input type="checkbox"/> 实时更新 <input type="checkbox"/> 日更新 <input type="checkbox"/> 周更新 <input type="checkbox"/> 月更新 <input checked="" type="checkbox"/> 年更新 <input type="checkbox"/> 其他 _____		
申请文件清单 <input checked="" type="checkbox"/> 登记申请书	<p>附加文件清单</p> <p><input checked="" type="checkbox"/>数据来源证明材料</p> <p><input checked="" type="checkbox"/>样例数据(供审核)</p> <p><input type="checkbox"/>数据样例附件(供展示)</p> <p><input checked="" type="checkbox"/>数据结构样例数据</p> <p><input checked="" type="checkbox"/>申请人承诺书</p> <p><input type="checkbox"/>数据安全合规评估报告</p> <p><input type="checkbox"/>其他证明文件</p> <p>_____</p>		
<p>全体申请人签字或者盖章</p> <p style="text-align: center;">钟正洋</p> <p style="text-align: right;">2026 年 2 月 24 日</p>			

