

基于无线传感网络测量的 PM_{2.5} 数据集

数据集介绍：本数据集是北京大学信息科学技术学院现代通信所在北京大学、西安电子科技大学**室内外**布置 PM2.5 无线传感网络后所收集的数据。目前的数据集提供 2018 年 3 月至 5 月传感器布置点的数据。另外，数据集生成时对一些不良数据进行了剔除。

数据集格式：

文件名	数据格式
basic_inf	设备 ID, 经度, 纬度, 修正系数, 是否在室外, 城市
beijing_inside	设备 ID, 采样时间, pm25, 校正后的 pm25, pm10, 校正后的 pm10
beijing_outside	
xian_inside	
xian_outside	

数据说明：

basic_inf 介绍了每个设备的基本情况，包括坐标，地点和 pm 系数的建议修正值。

- 经度、纬度只有室外检测设备有，记录在百度地图 BD09 坐标系中设备的位置。室内设备的该项属性为 None。
- 修正系数是将所有采样设备放在一起做校准后得出来的建议乘性系数。
Outside 属性为布尔类型，其值为 1 代表该设备在室外，其值为 0 代表该设备在室内。
- 城市有北京和西安两种取值。

beijing_outside/beijing_inside/xian-outside/xian-inside 分别记录北京/西安、室外/室内设备的采样数据。

- 属性 id 为设备号，可在 basic_inf 文件中查的该设备的相应信息。
- time 属性记录采样时间，采样频率为半小时一次。
- pm25、pm10 记录了原始采样数据，单位 $\mu\text{g}/\text{m}^3$ 。
- pm25_correction、pm10_correction 是原始数据乘以建议修正系数后的值，单位 $\mu\text{g}/\text{m}^3$ 。

数据示例：

```
basic_inf    {
    {
        "id": 124,
        "longitude": 116.316024,
        "latitude": 40.001547,
        "location": "未名湖",
        "correct_index": 1.145,
        "outside": 1,
        "city": "北京"
    },
    {
        "id": 125,
        "longitude": 108.84794,
        "latitude": 34.132813,
        "location": "校医院",
        "correct_index": 1.138,
        "outside": 1,
        "city": "西安"
    },
}
```

beijing_outside/beijing_inside/xian-outside/xian-inside

```
{
    {
        "id": 120,
        "time": "2018-05-04 12:00:00",
        "pm25": 37,
        "pm25_correction": 42,
        "pm10": 51,
        "pm10_correction": 57
    },
    {
        "id": 123,
        "time": "2018-05-04 12:00:00",
        "pm25": 45,
        "pm25_correction": 52,
        "pm10": 63,
        "pm10_correction": 72
    },
}
```