

数据说明：train_set.csv/test_set.csv

市长电话数据集是某北方城市政府部门冬季收到的“12345”服务热线的来电信息记录。数据集中的变量包括：

- (1) 单位名称：即每一条来电信息记录最终受理的政府部门；
- (2) 词汇出现的频数：已有词汇表中各个词汇在投诉建议的文本内容中出现的频数。

数据总共有 2000 行，每一行代表一条投诉建议中各个词汇的出现频数，第 2-6236 列的列名为某个词汇。