霍普金斯统计量，评估给定数据集是否存在有意义的可聚类的非随机结构。本文档用于记录各个表及其预处理后的Hopkins 统计量，该统计量计算距离时需要对数据进行标准化

|  |  |  |  |
| --- | --- | --- | --- |
| 表格名/数据名 | 预处理 | Hopkins statistic | 备注 |
| 三列100行均匀分布值 | 无 | 0.5±0.5 | 作为对照组 |
| 经营异常 | Z标准化 | 0.9999 | 聚类趋向明显 |
| 经营异常中三个指标单独 | 无/Z标准化 | 均大于0.999 | 单独聚类仍可成为有意义的标签  单独一个变量标准化与否均可 |
| 商标与信誉 | 无（均为计数量） | 0.977 | 聚类趋向明显 |
| 知识产权 | 无（均为计数量） | 0.9939 | 聚类趋向明显线程 |
| 知识产权四个指标单独 | 无（均为计数量） | 均大于0.99 | 单独聚类仍可成为有意义的标签 |

其余有统计意义的数据统计后为：

C:/Users/jnjga/AppData/Local/Programs/Python/Python37-32/python.exe c:/Users/jnjga/Desktop/university/服创/评价指标及代码/hop\_1.py administrative\_punishment.csv Hopkins=1.0

business\_risk\_rightpledge.csv Hopkins=1.0

business\_risk\_taxunpaid.csv Hopkins=0.9998549483504167

ent\_bid.csv Hopkins=0.9987688665445345

ent\_branch.csv Hopkins=1.0

ent\_investment.csv Hopkins=1.0

ent\_onlineshop.csv Hopkins=1.0

jn\_special\_new\_info.csv Hopkins=1.0

justice\_credit.csv Hopkins=1.0

product\_checkinfo\_connect.csv Hopkins=1.0

recruit\_qcwy.csv Hopkins=0.9970122605564818

recruit\_zhyc.csv Hopkins=0.9761778811395604

recruit\_zlzp.csv Hopkins=0.9997712613508607