k= 2

VRC方差率： 2918.09266318

轮廓系数： 0.93580493514

第0类数据占比: 98.8873435327%

第1类数据占比: 1.11265646732%

k= 3

VRC方差率： 3903.3059496

轮廓系数： 0.878358336217

第0类数据占比: 96.6156699119%

第1类数据占比: 3.01344459898%

第2类数据占比: 0.370885489105%

k= 4

VRC方差率： 4034.61988273

轮廓系数： 0.759384588412

第0类数据占比: 88.6416318962%

第1类数据占比: 0.370885489105%

第2类数据占比: 1.76170607325%

第3类数据占比: 9.22577654149%

k= 5

VRC方差率： 3907.47826281

轮廓系数： 0.723652100687

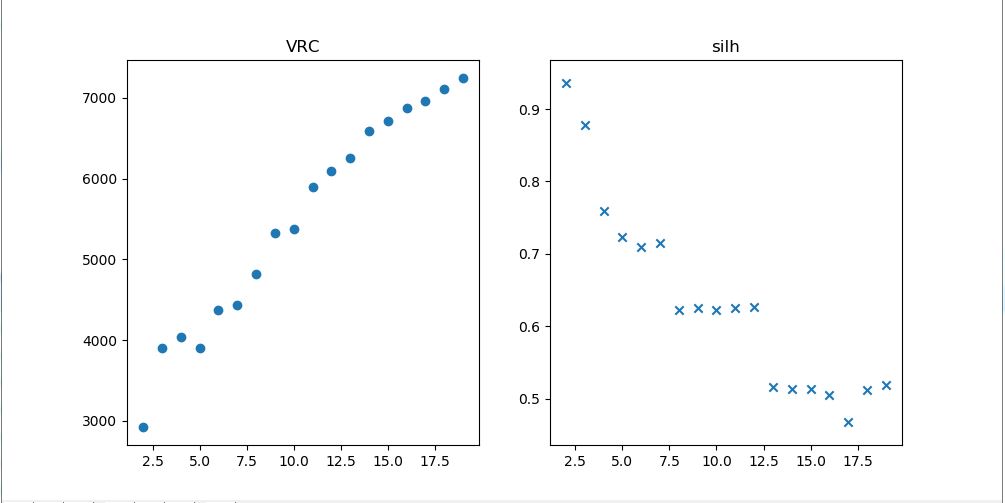
第0类数据占比: 85.3500231803%

第1类数据占比: 0.278164116829%

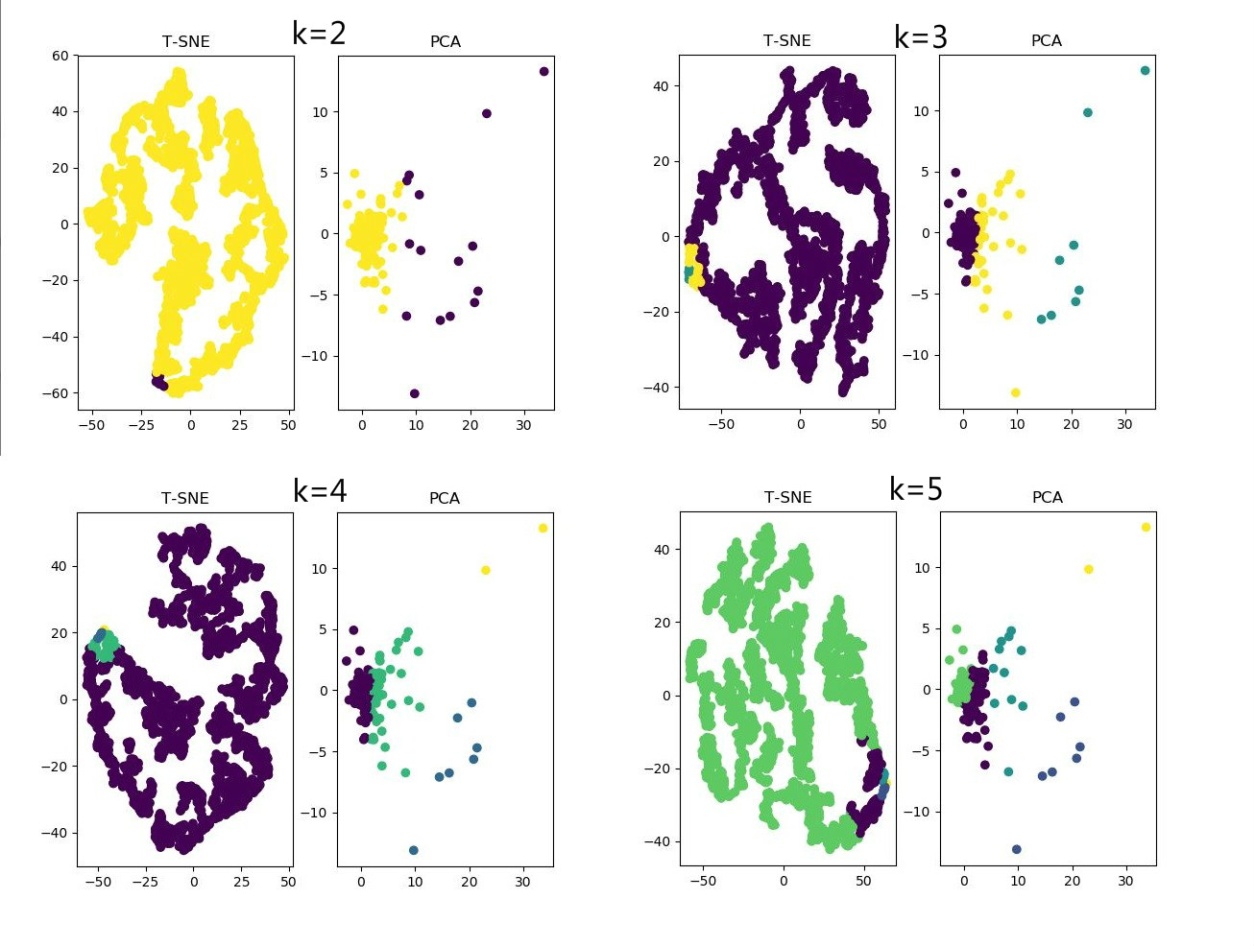
第2类数据占比: 2.50347705146%

第3类数据占比: 11.7756142791%

第4类数据占比: 0.0927213722763%



以资产负债表中：资产期末余额、资产年初余额、负债期末余额、负债年初余额、所有者权益期末总计、所有者权益年初总计 为特征进行聚类，随着值的递增，VRC方差率和silh轮廓系数的规律（上图，横轴k值）



图、T-SNE和PCA降维显示的分类结果

k= 2

VRC方差率： 2843.45265625

轮廓系数： 0.95845069548

第0类数据占比: 99.3045897079%

第1类数据占比: 0.695410292072%

k= 3

VRC方差率： 2397.77872037

轮廓系数： 0.89985124969

第0类数据占比: 97.9601298099%

第1类数据占比: 0.370885489105%

第2类数据占比: 1.66898470097%

k= 4

VRC方差率： 2577.87708922

轮廓系数： 0.87199162572

第0类数据占比: 97.0792767733%

第1类数据占比: 0.417246175243%

第2类数据占比: 0.0927213722763%

第3类数据占比: 2.41075567918%

k= 5

VRC方差率： 2730.88589801

轮廓系数： 0.802306518446

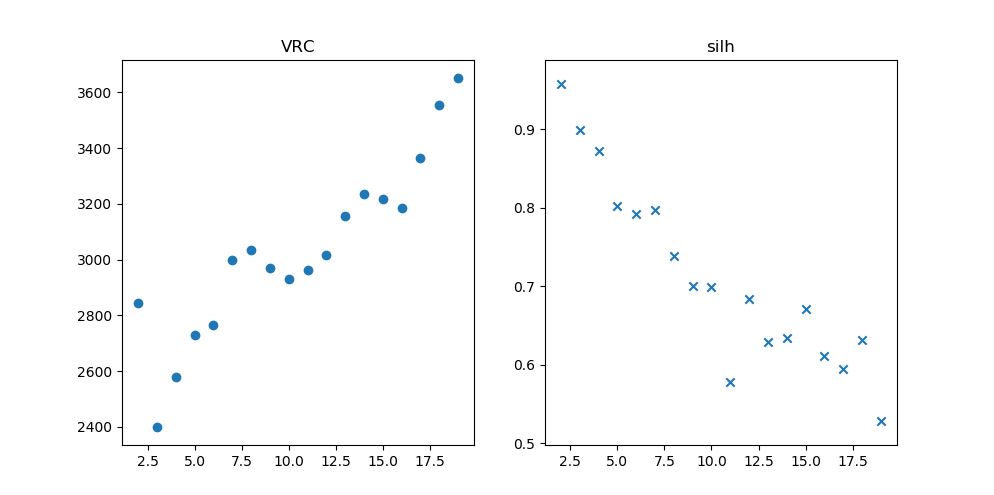
第0类数据占比: 93.2313398238%

第1类数据占比: 0.0927213722763%

第2类数据占比: 0.324524802967%

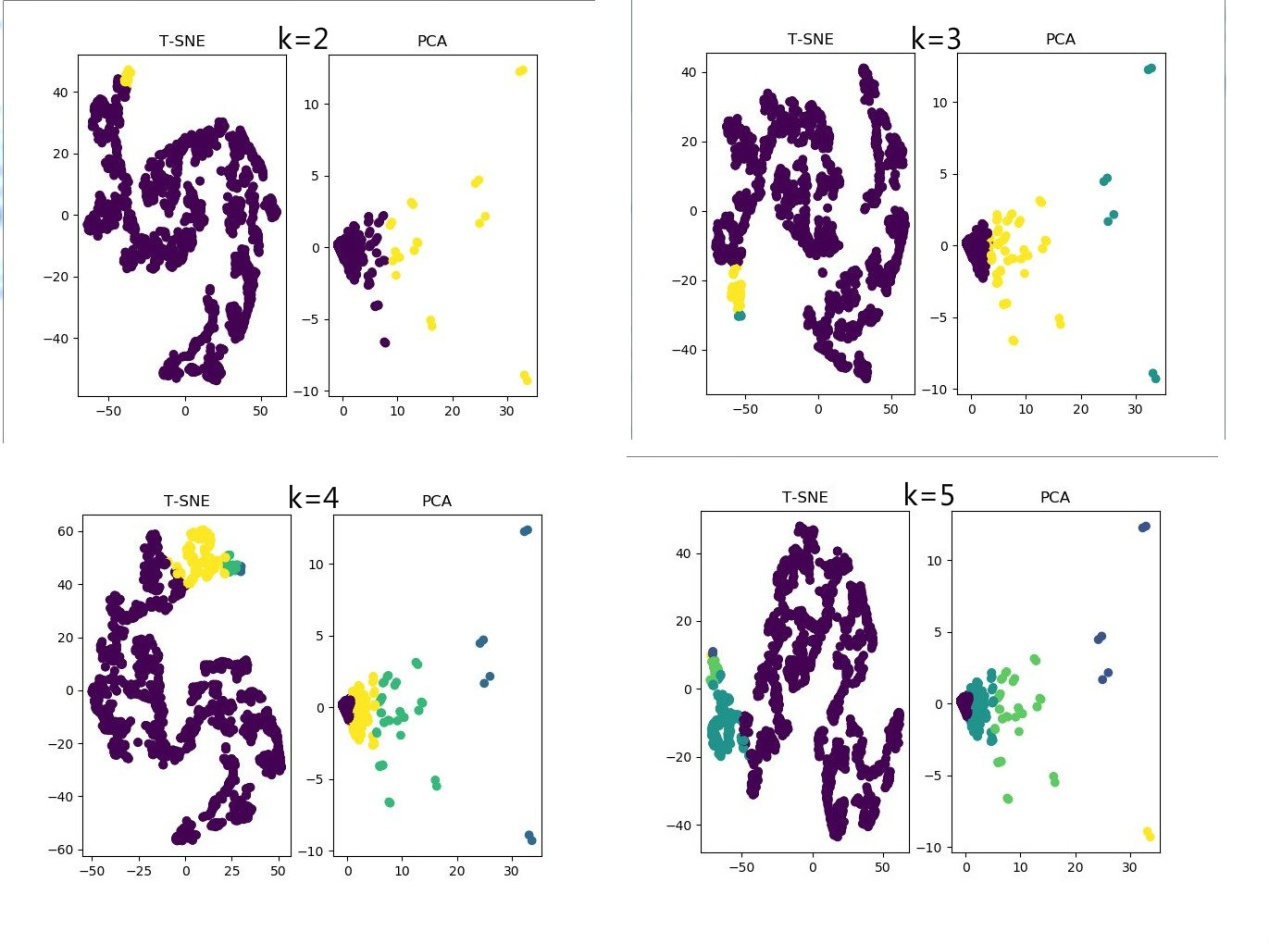
第3类数据占比: 0.50996754752%

第4类数据占比: 5.84144645341%



以资产负债表中：本期资产余额净增、本期负债余额净增、本期所有者权益净增

为特征进行聚类，随着值的递增，VRC方差率和silh轮廓系数的规律（上图，横轴k值）



图、T-SNE和PCA降维显示的分类结果

选取不同的特征进行聚类，根据silh轮廓系数的定义，k=2时为聚类最优结果

VRC方差率作为选取最优k的指标得出的结果很不理想