1. 了解一下<https://en.wikipedia.org/wiki/Record_linkage>
2. 安装recordlinkage

pip install numpy

pip install scipy

pip install pandas

pip install recordlinkage

1. 了解recordlinkage：

<https://recordlinkage.readthedocs.io/en/latest/notebooks/link_two_dataframes.html>

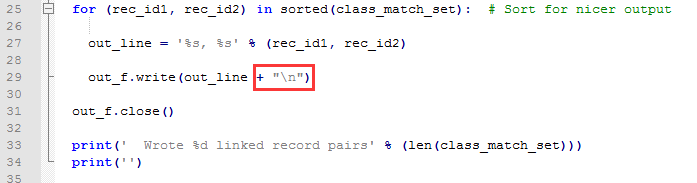
（建议有时间就看看，可以有更多方法去做，并且可以了解一下这个方法）

1. 代码，见附件



上图网址（<https://recordlinkage.readthedocs.io/en/latest/ref-compare.html>）

1. 对saveLinkResult.py修改



建议改回去。

1. 输出文件，目前是out.csv，可以自己改。注意输出有空格的，可以改一下输出的代码，去除空格