原始数据文件：

Company.xlsx 创业公司的基本信息

CompanyTeamRelation.xlsx 创业公司的团队信息

PeoplePositionRelation.xlsx 投资者所工作过的公司列表

代码文件：

* 网络中心度指标计算.ipynb
* util.py 涵盖一些用来处理的数据的函数

用到的函数主要有两个：

1. build\_dic\_function：生成字典用于快速查找
2. change\_year\_type：提取DealDate时间字符串中的年份

计算网络中心度的步骤：

1. 通过**build\_network()**的函数找到每一年之前，每个人对应的待过或者现任的公司列表，用到CompanyTeamRelation和PeoplePositionRelation两张表的数据；记录某人是否为teammember，用于之后的网络踢出无关人员的网络关系。
2. 通过**stage\_two()**函数将每年下，每个人所处公司列表之内的公司两两相连，即表示这两个公司通过这个人连接到一起，此处仅考虑创业人员（teammember）的关系网。其中calculate\_edge2()函数用于输入一个公司列表，返回一系列tuple表示两公司相连。
3. **stage\_three()**函数中遍历年份，在每一年中通过nx.Graph()创建图后，将stage\_two函数计算得到的网络边输入到图中，用nx.closeness\_centrality(G)命令计算得到每个公司的紧密中心度的值，并计算最大联通分支比例等辅助性数据。
4. 最后通过**save\_file()**函数将closeness的数据输出到excel中；若想将每年的节点总数、最大联通分支等信息直接输出到excel中，则运行最后一个block，将index\_save保存至excel即可。