附件中的txt是最新数据11.16

user.txt 主要包含总人数，性别，每周参与时长，用户身份和院系的统计，最后有关注邻接表（有向），之后用专门软件来画连接图

salon.txt 主要包含微沙龙月分布，每周各天累计微沙龙，发起地点分布，每场人数，发起时间，微沙龙领域分布

sign\_result 主要包含各院系参与人次，发起微沙龙次数，教师和教工的一些具体统计

coauthor\_result 主要包含微沙龙结交朋友数前100名单，埃尔德什数分析： 虽然埃尔德什数不大，但是有一半的人没有覆盖，所以用户联系的图中一定存在孤立的子图，可能需要进一步凝聚子群分析，也可以用深搜或广搜分出森林。

sex\_result 主要包含整体男女比例，微沙龙各领域男女数量和比例，各兴趣领域男女数量和比例，因为存在相当的人没有填性别，所以出现很多只有男生或只有女生情况。

interest\_result 主要包含各兴趣领域感兴趣人数，各兴趣领域微沙龙个数，兴趣领域邻接表（无向）

cross\_result 主要包含是否跨院系和本研交叉的各种分布，，比较多，比较绕

剩下三个分别时微沙龙名字，简介和感受的词频统计

缺少：

1. 时事的影响
2. 关注度有向图，兴趣领域连接图，用户连接图，都可以分析中心度和凝聚子群等等，需要专用软件来处理，邻接表都已经做好
3. 词云，需要再具体说一下怎么做
4. 另外之前教师的还有一个连接图，不明白其原理，而且我感觉做的意义不大。。其实很多连接图做出来因为点和边太多，显得太乱了，意义不是特别大。