#### **指导**

图是非常重要的一种数据结构，在社交媒体和现实生活中这一结构非常常见，且有大量分析统计基于图结构进行。本次作业提供了newmovies数据集（见资源/data/newmovies.txt），希望基于该数据，在程序中读取并存储用户节点信息，建立无向图结构，并进一步实现相关统计和可视化功能。

 1. newmovies.txt保存了相关数据，其中\*Vertices 34282 下的每一行为一个节点，表示一位明星、编剧或电影。每一行中属性以\t分割，分别为节点id，名称，节点权重，节点类型，其他信息（其他信息以";"分割)。注意，节点里的权重信息是原数据集提供的，本次作业用不到，另外edges部分的参数每行三个数，前两个是边所连接的节点id，第三个值均为1。

 2. 建立包GraphStat，实现相关网络的构建和可视化。其中

    1. 包Graph，用以实现点和图结构的创建，以及相关的基础统计功能

       1. 实现node.py模块

          1. 实现函数init\_node()，输入相关的节点文件，返回对应的节点列表，其中每个元素为一个节点

          2. 实现函数print\_node()，利用format函数，将节点属性输出至屏幕上

       2. 对graph.py模块，同上，同时实现图结构的序列化存储。

       3. 实现stat.py模块，基于图结构，进行基础的统计

          1. 计算图中的平均度并返回

          2. 统计图中节点某个属性的分布情况

    2. 包Visualization， 基于上述构建的图和节点结构，利用matplotlib等绘制相关的统计结果

       1. plotgraph.py

          绘制基于图的信息，如图的结构

       2. plotnodes.py

          绘制节点的属性分布，并提供结果保存

**结构示例**

GraphStat/

​            \_\_init\_\_.py

​            NetworkBuilder/

​                        \_\_init\_\_.py

​                        node.py

​                                    def init\_node(parmaters)

​                      返回字典，key为节点的属性，值为对应的属性值

​

​                                    def get\_xxx(node)...

​                                                获取对应的节点属性，如度等

​

​                                    def print\_node(node)

​                                                显示节点全部信息（利用format或者f函数）

​                        stat.py

​                                    def cal\_average\_dgree(graph)

​                                                计算网络中的平均度

​

​                                    def cal\_dgree\_distribution(graph)

                                                计算网络的度分布

​                        graph.py

​                                    def init\_graph(node\_list, edges)

​                             返回一个字典，分别存储节点信息和边信息

​                                    def save\_graph(graph):

​                                                序列化图信息

​                                    def load\_graph(graph):

​            Visualization/

​                        \_\_init\_\_.py

​                        plotgraph.py

​                                    def plotdgree\_distribution(graph)

​                                                度的分布图

​                        plotnodes.py

​                        绘制图中节点属性的统计结果

​                                    def plot\_nodes\_attr(graph,feature)

3. (附加)了解并学习networkx库，并对提供的数据集进行简单的分析和讨论。

4.（附加）了解并使用Gephi工具绘制网络结构。