

## 《程序设计思维》

### 2022年春季学期大作业选题

#### 大作业说明：

大作业是检验学期教学的重要方式，务必重视，其在期末总成绩中权重较高。大作业选题可根据自己兴趣确定，但必须报备获得任课老师认可。

大作业的评分关注重点包括：选题质量、选题难度、完成程度、代码注释、文案完备性以及界面美观等等诸多方面。文案包括：简要需求分析、简要项目说明、重点难点分析、简要测试数据、简要操作手册等等。

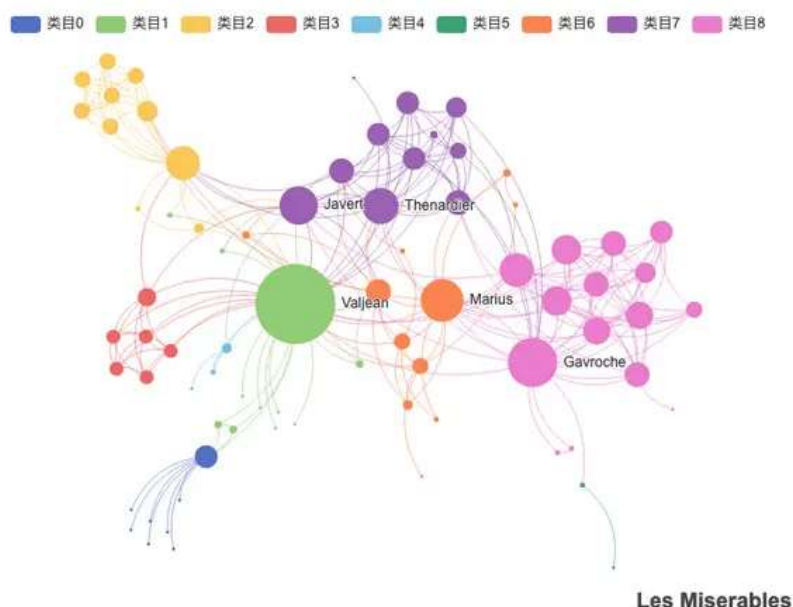
另外，大作业可以单人完成，也可以双人或三人完成组队完成。相对来讲，组队模式的选题难度要大于单人模式。

#### 备选题目：

1. **【可两人组队】**计算最佳新地铁线路。利用北京公共交通数据（地铁和地面公交）设计最佳北京新地铁线路。
2. **【可两人组队】**设计大站快线。在现有地铁数据的基础上，能否设计出大站快线，在特定时段为特定人流提供帮助，加快地铁系统效率；
3. **【可两人组队】**源代码语法正误检查，包括Python至少实现两门语言，结果呈现在Web界面上；
4. **【可三人组队】**对多个大陆高校排行榜数据综合利用，计算出新的排行榜，并详细说明新排行榜计分规则以及规则依据；
5. **【可两人组队】**升级LPOJ。下载LPOJ系统，对比学校的OJ系统，根据自己的理解优化升级该系统；
6. **【可两人或三人组队】**后台为Linux系统，编写Web界面，监控Linux服务器的执行状态，并适当控制，可利用第三方扩展包；
7. **【可两人组队】**利用openCV，校正手机或扫描的书籍页面图片，包括但不限于：梯形校正、背页（背后页面）阴影、污染乱点等等，要求过程详细

能在课堂讲解，能让没有图像处理背景的学生了解图像处理基础知识；

8. 【可两人组队】自动校正汉语拼音输入错误。无论是手机还是电脑执行汉语拼音输入时，总是可能有错误，比如：触碰到临近键位、前后鼻音等等。要求过程详细能在课堂讲解，能让相关知识的同学对此有所基本了解；
9. 利用网络多源数据，动态友好呈现历年中国各省GDP总量、进出口总值、人口总量及人口组成、婚姻、教育（含中小学学校、在读学生人数等）等数据的变化，并探索其相关关系；
10. 【可两人组队】换乘代价及换乘路径计算。利用作业提供的地铁数据，计算出换乘站换乘代价（换乘时间长短）。如：从4号线海淀黄庄站换乘到10号线海淀黄庄站的时长，以及10号线海淀黄庄站换乘到4号线海淀黄庄站的时长，其他站以此类推。在此基础上，计算出任意两站间的最合理路径；北京地铁网站（<https://www.bjsubway.com/station/zjgls/>）有站间公里数，也有相关时刻表（<https://www.bjsubway.com/e/action/ListInfo/?classid=39&ph=1>）等数据，可供参考；
11. 【可两人组队】《红楼梦》80+40研究。据考证，《红楼梦》前80回合和后40回合是不同作者，利用相关技术分析其异同，可利用jieba等第三方扩展包等。分析方法可包括：句子长短、字频分布、词语分布、虚词分布等。也可以将前80回合分为两个40回合，更利于比较。注意：jieba分词可能误判，要消除其误判等；
12. 【可两人组队】利用《三国演义》小说文本，分析人物关系，效果如下图



所示，建议采用eCharts相关可视化技术。可分章研究，利用jieba等等，获得《三国演义》人名数据，然后在每章内人物之间出现先后的距离【相隔多少文字，间隔文字越少越紧密】；

13. 利用提供的新浪新闻数据，发现词语。注意，不得利用jieba等第三方自然语言处理库。可考虑多个n-Gram组合利用。计算出的词汇，需要与jieba等分词库进行对比研究；
14. 【可两人组队】下载Harvard University的  
<https://projects.iq.harvard.edu/cbdb>的数据库，结合网络上各种文献，分析该数据库结构，各个字段（属性）含义。开发Web界面，可在地图上显示多人迁移路径，人际关系等多种应用；
15. 从美国约翰霍普金斯大学下载疫情数据，阅读相关文献，对疫情数据进行各种形式的可视化，可借助百度的ECharts和地图【可两人组队，建议多关注时间演进】；
16. 【可3人组队】对39所双一流高校网站评估。利用网络爬虫，对39所双一流高校（可减少数量，但不少于10所）官网进行多维度评估，包括但不限于：网页规范、图片规范、信息更新、栏目设置、栏目数量，热词跟踪等等；
17. 【可两人组队】SQLite3管理器。开发Web程序管理SQLite3数据库，包括：建立数据库文件、建立表、建立视图、建立索引、建立触发器等等。在此基础上，输入SQL语句，可在浏览器端查看执行结果等等。在此基础上，可提交将行列文本数据并将其中数据插入到数据库中；
18. 以ECharts某几类统计图为基础，开发可视化版参数配置。即在Web页面上，调整某些ECharts参数，及时获得显示效果，并能下载配置参数【根据规划功能确定人选数量，不超过3人】；
19. 在线Python程序设计。基于Web界面，可撰写Python代码，可运行程序并将程序运行结果反馈到界面。能智能判断变量名称撰写等错误或给予提示、包括代码提交前的静态检测以后提交后的检测等等【根据规划功能确定人选数量，不超过3人】；
20. 【可两人组队】根据Python代码（或C、Java、C++）生成流程图；