

《Python 程序设计》

——期末大作业选题

一、 要求：

- 1、 从以下题目列表中选一个进行设计开发
- 2、 以小组形式完成，**小组成员人数 2~3 人**，成员组合一旦成立，不允许随意变更!!! **每组一个题，不重复选题**

注意：请于 2022 年 10 月 23 日(第 8 周)前完成选题和组队。

3、 提交报告时间及内容要求：

- a) 每组需要提交一份**项目报告**（软件设计文档，格式参考“Python 大作业报告模板”），命名为**“题号_姓名_学号_项目报告.docx”**，对项目开展整体需求、功能和数据设计、部署方案、成果等方面的介绍，字数不少于 2000 字；
- b) 每人提交一份个人工作总结报告，命名为**“组号_姓名_学号_个人报告.docx”**，对个人在项目中主要负责的工作展开介绍和总结，要求内容真实、完整，逻辑清晰，格式规范；
- c) 提交**答辩 PPT** 一份，每个小组需要进行现场汇报（预计 2 次）：
 - 第 1 次汇报时间暂定：**2022 年 11 月 23 日晚 18:00~21:00**
 - 第 2 次汇报时间暂定：**2022 年 11 月 24 日晚 18:00~21:0**
- d) 提交**项目源代码及部署说明文档**（一个小组一份），打包命名：**“题号_组长名_源代码及部署文档.zip”**
- e) 所有文档**提交到头歌平台**（期末大作业），截止时间：**2022 年 11 月 22 日 23: 59 分**

二、大作业题目列表

1、“符号数学”（难度：★★☆☆☆）

要求：完成一个能求解各种数学计算的题库程序，解决家长、老师出题难的问题。

必备功能：交互界面，随机出题，算术符号功能设置（如加、减、乘、除、幂乘、位运算、开方、三角函数等），数值范围设置（如 10 以内，20 以内，100 以内等），题量设置，答题成绩报表分析

可选功能：错题集、答案解析、知识点分析、运算过程演示、报表打印等

2、“单词闯关”（难度：★★★☆☆）

要求：获取英文词汇表（文档导入或网络数据爬取），将词汇进行等级划分（如小学、初中、高中、考研、GRE 等），用 Python 编程实现单词选关挑战功能。

必备功能：关卡选择、关卡设置（如数量设置等）、游戏问答交互、随机出题、答题评分、错题本、学习进度分析

可选功能：可视化界面、1v1 对战、多人对战练习等

3、“网络行为预测”（难度：★★★☆☆）

要求：网络中的每一次点击、查看都隐含着用户的行为信息，要求应用 Python 语言，以用户在新闻 app 中推荐文章的点击、查看行为，分析挖掘用户的关注内容、行为习惯等信息，预测用户未来的可能行为。

必备功能：数据获取（具体 app、网页及数据源可以是自己拟定）、清洗、存储，（内容）分类、喜好预测、阅读习惯分析（时长、是否真实的关注信息）等

可选功能：历史关注内容可视化、行为习惯可视化、内容推荐等

4、“城市气候分析师”（难度：★★★★☆）

要求：近几年人口普查显示，我国城镇人口已突破 9 亿，城市资源及城市环境成为人们生产生活的主要关注对象，为打造绿色城市、宜居城市，需要对城市

气候环境进行提前预测分析，了解城市气候问题，提升城市可持续发展能力。

必备功能：①数据获取：以某一城市为例获取其近1年或3年内的主要城市气候数据指标（如气温、湿度、海拔、人口、绿化、科技、经济等）；②数据清洗、存储：将获得的气象数据进行组织、存储管理，实现简单的气候环境档案室；③数据分析可视化：如气候变化情况、人口迁移变化情况、绿化情况等。

可选功能：挖掘分析（气候与人口迁移关系、气候与经济发展关系、人才类别与科技发展或城市政策关系等等）

5、 “研途无忧”（难度：★★★☆☆）

要求：获取近3年全国各大高校录取专业、考研人数及录取人数、分数线等数据，进行升学情况分析

必备功能：数据获取（至少3所学校，2个相关专业）、清洗、图表可视化分析（如横向/纵向分析、生源和录取率的关系、人数和录取率的关系等）

可选功能：更多的数据分析，如高校排名变动与报名人数、录取率的关系，就业单位及就业率、薪酬等

6、 “学科前沿早发现”（难度：★★☆☆☆）

要求：科学研究要有前瞻性、探索性和创新性，为了更有效的挖掘发现学科发展的前沿趋势，奠定学科研究基础，需要对已有的研究成果、研究领域、问题等开展细致全面的挖掘分析。

必备功能：对给定数据集 arXiv，开展数据挖掘分析，包括但不限于研究方向统计、作者成果排名（如 Top10）、作者画像、学术关键词（学科趋势）等分析与可视化工作，得出合理、潜在的分析结论

可选功能：作者研究方向的迁移情况、论文作者关系模型、新论文类别分类等

7、 防抄袭神器（难度：★★☆☆☆）

要求：实现一个代码查重软件，检测多个代码文档，是否存在抄袭现象。

必备功能：一对一查重，重复率显示

可选功能：一对多查重，重复代码的标记，根据重复率筛选等

8、“偶像档案”（难度：★★★☆☆）

要求：结合爬虫技术，实现一个个人 idol 的管理网站，进行偶像信息的管理

必备功能：获取至少 3 类偶像（每一类各 5 人）的基本信息（姓名、出生年月、籍贯、类别、标签、作品、粉丝数等等）；数据分析及存储、数据分类展示等

可选功能：入粉时间、偶像作品特征、理由等因素之间的关系、不同类别偶像间属性的关联度分析等等挖掘用户特征、偶像推荐（好友推荐）等

9、“智能知识侦查助手”（难度：★★★★☆）

要求：爬取豆瓣图书 Top200 的书籍，编写对图书信息侦查的问卷生成程序，如进行已读/未读/没听过等选项的调查。

必备功能：数据的获取、清洗存储；问卷模板制作（如判断、选择等题型）；自动问卷的生成（如书籍的随机抽取、顺序的随机排布等）；结果数据分析统计等

可选功能：可视化界面，书籍知识点领域分类，个人认知等级评估（如根据结果进行书籍的个人认知分析）等

10、“诗词竞答”（难度：★★☆☆☆）

要求：获取存储唐诗、宋词等约 300 篇，实现一个随机诗词竞答的 python 程序应用

必备功能：随机组题（如 10 题一套，10 分一题；20 一套则 5 分一题等），随机抽取诗词的其中一句作为题目，进行单选答题（诗句所属题目、作者、上一句、下一句等），统计总分。

可选功能：错题集、多次结果统计可视化等

11、“诗词创作之路”（难度：★★★☆☆）

要求：每一首华美的诗篇背后总有一段曲折的创作之路，如杜甫的《三吏三

别》，李白的《蜀道难》，毛泽东的《七律·长征》等等，设计实现一个刻画诗词赏析及诗人画像背景的程序应用。

必备功能：采集获取唐诗、宋词或近现代诗词等约 300 篇及其背景信息，提供按作者、诗词名、年代、类别等的搜索功能，并展示诗人诗词创作的背景故事，绘制作者的个人生平画像

可选功能：作者间的关系图谱等

12、 “美图秀秀”（难度：★★☆☆☆）

要求：编写程序，实现多个图片的拼接整理功能

必备功能：图片选择、拼接模板设计管理、2-6 张图片的拼接等

可选功能：自由拖动调整、抠图等

13、 邮件管理器（难度：★★☆☆☆）

要求：编写程序，实现电子邮件的撰写、群发功能，邮件中支持附带图片和文件

必备功能：邮箱地址管理、邮件收发、定时发送等

可选功能：群发单显、撤销等

14、 “杭州楼市”（难度：★★★☆☆）

要求：编写程序，爬取近 3 年杭州热门楼盘的房价数据，实现对房价数据的分析处理。

必备要求：数据爬取、清洗功能，房价、楼盘、位置、年度的相关性分析和可视化展示等

可选功能：热门区域的热力图展示、房价预测、入住率、二手房买卖情况等

15、 在线选题系统（难度：★★★★☆）

要求：编写程序，实现一个在线选题系统，便于集中在线选题。

必备要求：发布题目、设置规则（如设置小组、人数、单选、截止时间等）、

实时选题、选题成组确认等

可选功能：题目批量导入、模板导入、富文本编辑、打印导出等

16、 领域会议（难度：★★★☆☆）

要求：每个专业（学科）领域都有各自公认的、顶级的会议及组织，为有效掌握各学科技术的发展动态，实现一个及时获取跟踪计算机相关学科领域的各大会议组织及时间信息的应用程序。

必备要求：计算机相关专业领域设置；领域相关会议信息获取（如会议名、会议时间、地点、主办方、投稿时间、截止时间、会议主题等）；年度会议时间线（实时状态）；主题画像（主办方组织会议画像、每年会议主题变迁画像等）等

可选功能：与会人员关系图谱、年度会议关键词等

17、 “随想绘图板”（难度：★★★★☆）

要求：想象力和创意思维往往需要通过文字或图的方式进行表达展示，为激发孩子的想象力、培养创造力和执行力，用 Python 程序开发一个有利于孩子创意制作的随想绘图板。

必备要求：颜色、粗细、笔或几何图形等基本信息的选择；类似画图工具实现鼠标点击移动后的作品绘制；颜色填充、作品保存等

可选功能：模板部件的创建、导入及使用等、见图猜字，看字画图等

18、 拼图游戏（难度：★★★☆☆）

要求：拼图游戏是一款经典的智力小游戏，通过将一副完整的图片切割成大小相等的 9、16 或 $N \times N$ 块，随机打乱并去除其中某一块小片后，通过左右上下移动完成图片的复原。

必备要求：图片选择加载；游戏生成；游戏关卡；得分统计

可选功能：1v1 对战等

19、 记账/报账小能手（难度：★★★☆☆）

要求：线上消费已成为生活的主要方式，通过支付宝、京东、淘宝、微信等app 每天、每月都会产生许多账单，多样的消费方式提供了生活的便捷性，但同时让个人理财/记账信息变得更加繁杂，不论用于生活记账还是办公财务报账，发票的管理、开销的记录都难以清晰的整理归档。编写一个程序，实现一款灵活记账、发票归档等功能的个人理财神器。

必备要求：账单数据的增删改查（月、季度、年账单的查询）、额度的管理、超额预警提醒等等，账单消费类别的统计、分析，购物方式/行为的画像分析等

可选功能：个人消费账单数据的导入，发票的上传/下载/打印/识别归档等

20、 数据智能标注处理器（难度：★★★★☆）

要求：近年来，深度学习、神经网络、人工智能等概念/技术的发展如火如荼，在交通、医疗、生态、农业、服装等领域取得了突破性进展，智慧交通、远程医疗、智能农业、虚拟试衣等场景陆续成为可能。但每种技术、方法的推出离不开前期数据的准备、实验和演算，数据是技术应用发展的基石。编写一个程序，实现对数据的标注处理，如图像分类标记、目标检测标注、图像分隔标注等。

必备要求：对图像或文本数据进行数据标注操作，包含但不仅限于数据的导入、数据的清洗、数据的标注、数据的管理等

可选功能：开展智能标注，通过提供已标注数据，进行预学习，对新样本数据进行标注；数据的归档管理、数据的标准化等

21、 职业规划师（难度：★★★★★）

要求：大学是从学生过渡到社会的桥梁阶段，在大学，你开始接触专业，开始参与社团，开始安排自己的日常作息，开始规划未来，那么未来究竟如何？一千个读者有一千个哈姆雷特，每个人的生活定位都有自我的意义，如何实现自我价值都将体现在职业的规划中，那么如何规划职业呢？有哪一些职业？请设计一款程序，通过信息化、数字化带给我们的便利性，实现个人职业规划的功能。

必备要求：调研设计职业规划相关的因素/条件，采集数据，分析数据

可选功能：职业/行业的领军人物、特点、职业的要求、各职业的关系等等

22、 课程故事会（难度：★★★★★）

要求：每一门课程的开设都有其历史背景和社会意义，如 Python，因教学应用出现，因技术发展蓄力，因机器智能爆发。每一门课程也都有它自己的知识脉络，讲好一门课或学好一门课则需要对其知识脉络、知识含义有深入浅出的认识，以自己已学的课程为基础，编写程序，讲好课程故事。

必备要求：至少 1 门课，包括课程信息的录入/导入（课程类别、名称、知识点、知识点的背景/涵义等）、课程知识导图绘制、输出

可选功能：2 门课及以上，呈现课程知识点间的衔接、关联，课程知识点的应用领域、场景等

23、 “资产管理”平台（难度：★★★★★）

要求：大数据时代，数据已成为一种无形的资产。信息处理的基础是数据，业务开展的基础也是数据，如何采集并管理好数据成为当前社会及行业发展的一个热点问题。在互联网中，数据各式各样，可能是纯文本、可能是图片、声音、动画等等，如何更便捷、高效的管理这些数据也是大数据技术的一个研究热点。请编写程序，实现自己感兴趣领域的数据资产管理平台，实现对数据信息的采集、管理和分析功能。

必备要求：至少选择 1 个领域大类（如环境、气候、交通、教育、经济等），采集获取该领域相关的数据（以教育为例，涉及国家或省市教育相关的制度文件、办学资源情况、教育水平情况、科研情况、心理健康情况等等），数据的描述/应用场景等，实现数据的信息检索、统计、分析等等

可选功能：对数据开展潜在的信息挖掘分析、市场/发展等的关联分析/决策分析