

2021 年春季《大学计算机基础》(理科) 大作业

题目要求

一、要求和说明

从下面给出的 5 道题目中, 选择 1 道题目, 采用 Python 语言实现程序并撰写实验报告。要求独立完成。

(1) 在实验报告中, 需要介绍作业完成功能、实现方案、设计思路 (**给出关键代码**) 和涉及的知识点、创新点或增加功能、遇到的难点及解决办法 (**给出关键代码**)、GUI 初始界面、程序运行结果截图等。

(2) 进行课程学习总结: 课程收获、课程难点分析、教师授课评价、助教评价、课程进一步改进建议。

二、关于申优答辩

大作业设有申优答辩环节, **必须参加申优答辩大作业才可能获得优秀** (90 分及以上)。如果你认为自己的这次作业完成得很好, 可以申请优秀, 则必须参加答辩, 答辩形式由各班商讨决定。答辩现场表现占大作业成绩的 40%, 另 60%根据源程序完成情况、实验报告质量评定。大作业成绩占总成绩的 15%。

说明: 你也可以不参加申优答辩, 则大作业成绩最多只能是良 (89 分), 主要根据源程序完成情况、实验报告质量评定。

申优答辩具体时间由教师、助教及全体同学以小班为基本单位商议决定), 答辩顺序由任课教师提前排定。每位同学有 4 分钟的展示时间以及 1 分钟的回答提问时间。

在任课教师规定的提交截止时间之前, 将最终版源代码、实验报告 (包括课程学习总结)、相关材料 (如程序运行必须的文件、图片, 或者程序运行所必须的其他模块及其安装说明)、答辩用幻灯片 (参加申优答辩者) 放入以“学号+姓名”命名的文件夹中, 压缩后提交至课程中心指定栏目。

大作业不允许抄袭! 我们将对作业进行查重。一旦发现抄袭, 则抄袭者和被抄袭者均为 0 分!

题目 1：本地简单数据分析平台

一、题目描述

当今社会中统计学、大数据分析等已经成为无论文理的各个学科中必要的研究途径以及展示数据最好的形式，其便捷性及强大的功能也是各个软件所追求的目标之一。

选择本题，你需要完成一个程序，包含一个 GUI 的界面，能够清晰地引导用户导入数据文件，并且给出至少两种数据分析及展示方法，最后能够将结果导出为用户指定的格式。

二、基本实验要求

本题你必须给出一个 GUI 界面，需要清晰直观地给出导入和导出文件的按钮，在导入文件之后能够给出数种可以进行一键分析的按钮。

根据用户选择的分析方法，需要自动检测导入的数据是否符合要求（例如单变量数据不能够进行相关分析），并给出相应的提醒或者是报错信息。

展示方法中可以由用户自定义横纵坐标刻度、标题以及图例外观等相关参数，支持双变量做图等功能，导出前可以由弹窗等形式进行预览。

三、必做部分

你的程序必须有一个 GUI 页面，要求简洁美观，能够一键完成用户所需的操作：

（1）导入数据文件。基础的单变量数据文件规定为一个 $n+1$ 行两列的 csv 文件。

其中包含 n 个数据，例如下表：

年份	车辆总计（万辆）
2008	930.61
2009	1087.35
2010	1133.32
2011	1263.75

要求 n 不小于 100，数据的来源可以根据自己的专业喜好来选择，可以但不限于从以下网站获取：

a) 国家统计局: <http://data.stats.gov.cn/index.htm>

b) Science Data Bank: <http://www.sciencedb.cn/index>

导入的方式可以是直接从当前目录读取特定名称的 `csv` 文件，也可以给出一个文本框，由用户输入文件名。

(2) 开展数据分析，必须能够至少输出包括最基本的描述统计量（平均数，中位数，众数等）。

在导入数据后，应给出相应的数据分析按钮，用户点击后应显示出分析结果。

(3) 对于描述统计结果可视化，必须包括以下两种方式：

a) 柱状图

b) 折线图

(4) 将 `Matplotlib` 可视化结果可以用弹窗进行预览，并且导出保存，图片格式为 `png` 格式。保存的位置任意。

四、选做部分

以下内容非必做项，每完成一条选做项，都可以给大作业带来额外的加分，你同样可以用其他额外的内容来丰富你的大作业，以获得加分。以下为提供的一些思路 and 方向。

(1) 在导入数据方面，支持更多的文件格式和数据文件内容，比如 `json`、`xls` 文件，多变量数据等。

(2) 在数据分析方面，添加更多的方法，可以包括但不限于：

a) 广义线性模型

b) 非参数检验

(3) 在做图方面，支持更多的做图类型，例如饼图等，支持用户自己定义图例，图线颜色、标题等。对于多变量的数据，可以进行多线图等多变量的图形绘制。

(4) 在 `GUI` 方面，能够将导入的数据使用表格进行展示。

(5) 在用户选择了分析方法后，可以自动检测导入的数据是否符合要求（例如单变量数据不能够进行相关分析），并给出相应的提醒或者是报错信息。

(6) 导出时，你可以利用 `Markdown` 文本，将数据与分析结果进行整合，最后生成一份详细的数据报告。

以上选做部分不要求全部完成，也不必局限于给出的这些内容。

五、实验指导

1、`Pandas`：强大的分析结构化数据的工具集对于一般的二进制文件当然可以用 `open()` 函数直接实现，但是对于 `json`、`csv` 等需要二次处理的文件，或者是 `xls` 等

非二进制文件，可以使用 **Pandas** 进行处理，同样 **Pandas** 也可以写入这些文本。
在数据处理时 **Pandas** 的 **Dataframe** 也能使整个工作流更简化。

官方中文网站: <https://www.py pandas.cn/>

2、NumPy: 更强大的科学计算的基础软件包

在处理更大量的数据和多维数组时，利用 **NumPy** 能更高效和简便地处理数据。

官方中文网站: <https://www.numpy.org.cn/>

3、Matplotlib: 基础的绘图工具

这个不做过多的说明，课程内的内容。

官方中文网站: <https://www.matplotlib.org.cn/>

4、PyQt5 | Tkinter: GUI 开发库

使用 **Tkinter** 或者 **PyQt5** 完成大作业均可。非官方中文教程:

PyQt5: https://maicss.gitbooks.io/pyqt5/content/hello_world.html Tkinter:

<https://www.runoob.com/python/python-gui-tkinter.html>

题目 2：数值计算

一、 题目描述

随着大数据时代的到来，人们需要处理的数据量也指数式上升，数据科学的基础是优化，而优化的基础是数值计算。

选择本题，你需要写一个程序实现数值计算中的若干个算法，并编写一份文档能够清晰地引导他人使用这些算法。

二、 基本实验要求

本题你仅可以使用 `math`, `numpy` 两个包实现以下内容：

算法目的	算法名称	具体功能
解非线性方程（组）	Newton 法	可以根据不同的精度要求得到近似解，调用函数时可以选择是否输出中间迭代值。
得到正交化多项式组		对于给定离散的有限个点与一正整数 n , 求此离散点集上的正交多项式系 $\{p_i(x)\}_{0 \leq i \leq n}$ 。 对于给定的正整数 n , 求 legendre 多项式系 $\{l_i(x)\}_{0 \leq i \leq n}$ 。
求解积分	Gauss-Legendre 方法	对于给定的函数与区间，求此函数在此区间上的积分近似解。

三、 必做部分

上述每种算法都要以函数的形式呈现，但是一种算法可以由多个函数组成。此外，你还需要写一份详细文档说明各个函数传入的参数与返回的内容，以及用流程图、伪代码等形式给出算法的详细描述。

此外，请在实验报告中给出以下运行结果，绘制函数图像，并标出近似解的位置：

- (1) $x - \ln x = 2$ 在 $(2, \infty)$ 内的近似解与中间的迭代值。
- (2) $\{1, 2, 3, 4\}$ 上的正交多项式系 $\{p_i(x)\}_{0 \leq i \leq 3}$ 。
- (3) legendre 多项式系 $\{l_i(x)\}_{0 \leq i \leq 3}$ 。
- (4) $\int_{-1}^1 \sqrt{x+1} dx$ 的近似值。

四、 选做部分

以下内容非必做项，每完成一条选做项，都可以给大作业带来额外的加分，你同样可以用其他额外的内容来丰富你的大作业，以获得加分。以下为提供的一些思路 and 方向。

- (1) 利用其他数值算法计算上述三个方向，例如利用拟牛顿法求解非线性方程组、利用 **Newton-Cotes** 方法求积分。

- (2) 实现其他数值计算关注的方向，例如拟合曲线、拟合曲面。
- (3) 增加函数灵活性，例如通过参数来控制函数是否输出一些重要中间值。
- (4) 扩展文档，例如加上每个方法的适用条件与优势，加上每个方法的样例。
- (5) 以上选做部分不要求全部完成。

五、实验指导

1. NumPy: NumPy 中内置许多线性代数包可供使用。官方网站为: <https://www.numpy.org.cn/>
2. 建议在开始写每个算法的代码前，先将流程图或伪代码写出来，使思路更清晰，也减少 Debug 时间。
3. 算法具体内容可以网上查询，也可以参考有关书籍，例如《数值分析》，朱晓临。

题目 3：空气质量信息聚合平台

一、设题目的

1. 强化课上学习的 GUI 相关知识，能够使用 Python 设计 GUI。
2. 了解地理信息系统知识，并掌握如何用 Python 进行地理信息系统相关操作。
3. 了解时间序列知识，能够使用 Python 相关库对时间序列数据进行分析与预测。
4. 学会自主阅读 Python 第三方库文档。
5. 强化将大规模任务分解成若干部分的小任务意识。

二、题目背景

党的十八大以来，习近平总书记多次在不同场合就打赢蓝天保卫战等工作发表重要讲话。“环境就是民生，青山就是美丽，蓝天也是幸福”，“基本消除重污染天气，还老百姓蓝天白云、繁星闪烁”，这些都是习近平总书记对打赢蓝天保卫战的重要指示。[1] 蓝天越来越多，也是百姓对改善空气质量的殷切期盼。

三、题目要求

本题希望你整合一个地区各地的空气质量信息，结合地理信息系统的知识，制作一个 GUI，将数据呈现出来。如果有条件的话，希望你可以根据不同地区的不同空气质量特征对空气质量数据进行分类，另外，也希望你可以对今后一段时间各地的空气质量进行预测。

本题所指空气质量指 AQI（空气质量指数）及《环境空气质量标准》（GB 3095-2012）规定的六项基本污染物（二氧化硫、二氧化氮、臭氧、一氧化碳、PM 10、PM 2.5）浓度。

如果你要绘制其他地区（如欧盟、美国等）的空气质量地图，请优先考虑当地组织规定（欧盟：欧盟环境总司，美国：美国环保署 USEPA）的首要污染物（或优先控制污染物）。

四、必做项

1. 根据你收集到的空气污染情况的表征参数（如 AQI、SO₂、PM 10 浓度等），选择你感兴趣或关心的五个以上的城市，绘制包含其 AQI 具体表征参数的直方图。
2. 绘制最近五天内日均的空气质量地图。（日均 AQI 的标准详见 GB 3095-2012）

五、选做项

1. 将这些图片放在 GUI 中进行显示，可选择用桌面端 GUI（Tkinter, PyQt）或网页端 GUI（Flask, Django），可提供交互查询的功能。

2. 实时获取空气质量数据并绘制出实时空气质量地图。（注意：实时报 AQI 和日均 AQI 的标准不同，详见 GB 3095-2012）
3. 能够对空气质量进行预测，进一步地，可以对你的预测效果进行解释。
4. 当然你也可以锦上添花，不局限于这些选做项。

六、说明

1. 非常重要!!! 在绘制中国地图之前，请先学习自然资源部的[国家版图地图知识](#)。国家版图绘制错误是原则性错误，如犯此错误，大作业直接按不及格处理。（如南海诸岛、钓鱼岛、赤尾屿或台湾省等不能绘制在地图上，务必特殊说明！）
2. 考虑到地图绘制时间可能较长，你可以预先绘制好并保存在本地，但是你需要解释绘制过程。

七、题目指导

所需库

1. Python GUI 绘制工具：推荐使用基于 Qt 框架的 GUI 库，比如 PyQt5 等，其他的还有 wxpython、Tkinter 等。
2. Python 地理信息系统及可视化工具：推荐使用 geopandas（操作类似 pandas）和 geoplot，其他的工具还有 mpl_toolkit、pycharts 等。
3. 获取互联网数据可能会用到：urllib3、requests、BeautifulSoup 和 json。

数据来源

不同地区的空气质量指数不尽相同，参见维基百科 Air Quality Index。

可以使用 API 进行数据获取，推荐使用聚合数据 www.juhe.cn。

Jim 提供中国地图的 .geojson 文件，数据来源于 GeoAtlas

<http://datav.aliyun.com/tools/atlas/>；其他地区的地图文件请自行获取。

题目 4: Python-based Photoshop

一、题目背景

kk 经常使用 Photoshop 处理图像，但他的电脑运行十分缓慢，所以他想要自己实现一个具有 Photoshop 基本功能的程序，来方便快捷地对他的图片进行操作。

二、题目要求

在本题，你需要实现一个具有图像处理功能的 Python 程序，示例界面如下所示，主要需要包括菜单栏、工具区、画布等区域（如果实现了其他功能，可以加入相应区域）。

要求该程序能够打开或创建图片，并包含一系列基本的图像处理工具，例如画笔工具、橡皮工具、选区与移动工具、裁剪工具等。

三、必做项

你的程序应当能够打开图片，使用以下几个工具（**任选三个完成**）对其进行编辑，并能够保存编辑后的图片。

- 1. 选区工具：**在图片上某个位置点击鼠标，并拖动鼠标，可以获得一个矩形选框（你也可以选择其他的选区形状，例如圆形）。
- 2. 移动工具：**如果当前不存在某个选区，则可以通过点击与拖动鼠标，对整个图片进行移动；如果存在某个选区，则可以对选区内的图像进行移动。
- 3. 画笔工具：**在图片上点击与拖动鼠标，自由地改变图像中的像素点。（你可以只实现固定颜色、固定粗细的画笔）
- 4. 橡皮工具：**在图片上点击与拖动鼠标，自由地清空图像中的像素点。（你可以只实现固定大小的橡皮工具）
- 5. 裁剪工具：**在图片上点击与拖动鼠标，可以裁剪出图像上的一个矩形区域。
- 6. 模糊工具：**在图片上点击与拖动鼠标，可以将图片上的一片区域模糊。（你可以自己选择模糊算法）

四、选做项

在必做内容的基础上，你可以选择性地实现以下几个功能。

- 1. 自由选区工具：**可以选出图片中的不规则区域。你可以自己来设计如何选出这个区域，例如用鼠标在图片中画出一个不规则区域，并将这个不规则区域作为选区；或者用鼠标在图片上选中几个点，并在敲击回车后根据这几个点来建立一个多边形的选区。
- 2. 可自定义的画笔工具：**你可以升级必做内容中的画笔工具，例如，可以调整画笔的粗细、颜色（可以使用取色笔在图片中取色，或者在调色板中选取某种颜色）、硬度（在画笔的边缘处过渡到透明）。

3. **可自定义的橡皮工具：**同上，可以调整粗细、硬度等。
4. **自由变换工具：**可以对选区进行旋转、镜像、切变、扭曲、透视等变换的工具。
5. **图层功能：**通过建立图层，你可以对图片的不同层进行编辑，在编辑其中一个图层时，其他图层不受影响。
6. **其他：**你可以为你的图像处理工具补充一些其他具有一定实用性的功能。

五、题目指导

1. 本任务是一个开放性任务，你可以在给定的框架下，自己设计每个工具具体的行为。
2. 为了完成这项任务，你可能需要学习以下几个 Python 包的用法：

Tkinter: <https://docs.python.org/zh-cn/3/library/tkinter.html>

Pyqt: <https://www.riverbankcomputing.com/static/Docs/PyQt5/>

python-opencv: https://docs.opencv.org/3.0-beta/doc/py_tutorials/py_tutorials.html

pillow: <https://pillow.readthedocs.io/en/stable/index.html>

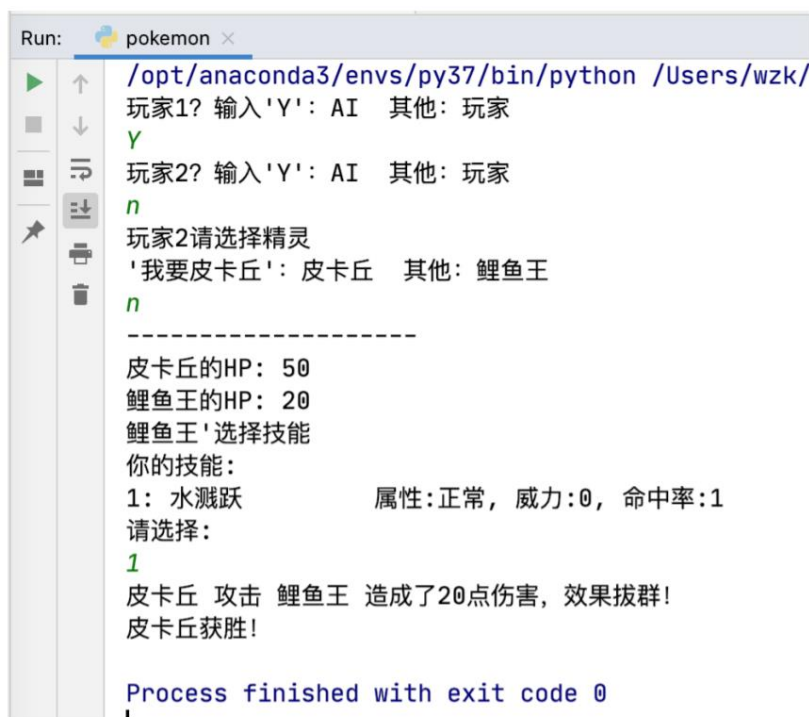
题目 5 口袋妖怪

一、题目描述

口袋妖怪（又名精灵宝可梦、神奇宝贝）是 nyima 最爱玩的游戏。他想编写一个简单的命令程序来模拟口袋妖怪的对战场面，这样他就可以与电脑或者好盆友对战了。

二、基本实验要求

你需要实现一个精灵对战的命令行小游戏，示例如下所示（仅展示了对战页面，绿色为控制台输入）。双方选取所用的精灵后，轮流选择技能进行对战。



```
Run: /opt/anaconda3/envs/py37/bin/python /Users/wzk/
玩家1? 输入 'Y': AI 其他: 玩家
Y
玩家2? 输入 'Y': AI 其他: 玩家
n
玩家2请选择精灵
'我要皮卡丘': 皮卡丘 其他: 鲤鱼王
n
-----
皮卡丘的HP: 50
鲤鱼王的HP: 20
鲤鱼王'选择技能
你的技能:
1: 水溅跃 属性:正常, 威力:0, 命中率:1
请选择:
1
皮卡丘 攻击 鲤鱼王 造成了20点伤害, 效果拔群!
皮卡丘获胜!

Process finished with exit code 0
```

其中，对战规则如下：

1. 每回合双方分别决定其精灵要使用的技能（每个精灵最多有 4 个技能），速度快的精灵先使用技能，速度慢的后使用技能。
2. 技能大致分为伤害类技能和变化技能两类。伤害类技能可以减少对方的 HP，需要根据属性相克关系、技能威力、攻击者的攻击力和被攻击者的防御力计算伤害；变化类技能可以改变双方的攻击力/防御力，或者给对方施加麻痹、中毒、睡眠等异常状态（参见参考资料）。
3. 一方 HP 为 0 时结束游戏，另一方取得胜利。

更详细的资料可以参考 [神宝百科](#)

三、评分细则及加分项

1. 必做部分

1. 屏幕交互：至少应包括

- 开始时：输出自己的[学号姓名](#)，根据用户输入决定双方是玩家还是 AI，显示玩家可选的精灵和技能并进行选择
- 对战时：每一轮需要显示对战双方的当前 HP、最大 HP、异常状态、技能列表和技能详细信息；双方轮流从输出的技能列表中选择要使用的技能；双方均选择完后，速度快的精灵先使用技能，速度慢的后使用技能
 - 技能详细信息：显示技能的 PP（即最大使用次数和剩余使用次数）、属性、威力、命中
- 游戏结束：判定获胜的一方并输出

2. 精灵选项：从[宝可梦列表](#)选取。至少应包括 5 只精灵，这些精灵应包括至少 5 个属性（因为可以有精灵为双属性），可以在具体精灵页面查看(举例：[皮卡丘](#))

3. 技能选项：从[技能列表](#)中选取。至少应包括 20 个技能，其中必须有伤害类技能；每个技能包括属性、PP、威力、命中率等内容

4. 伤害计算：参见[招式伤害计算公式](#)

- 计算公式：
$$\text{伤害} = \left(\frac{2 \times \text{等级} + 10}{250} \times \frac{\text{攻击}}{\text{防御}} \times \text{威力} + 2 \right) \times \text{加成}$$

5. 实际的宝可梦游戏中伤害计算需要考虑双方的等级、攻击、防御等，你可以预先统一每个精灵的攻击、防御、等级以便进行计算，也可以根据选做内容中的种族值进行设定

- 加成：中要求[属性相克加成](#)（参见[属性克制](#)一节）、[击中要害](#)（有一定概率使伤害乘 2）、[属性一致加成](#)（精灵属性和技能属性一致时伤害乘 1.5）必须考虑，其余可选

6. 简单的 AI：为了方便人机对战，要求实现简单的 AI，至少应实现：选择伤害数学期望最高的技能（考虑属性加成、技能威力、命中率等）

7. 属性克制：大家一定还记着期中时的 nyima 的口袋妖怪吧，这里将会用到那道题中需要从[属性相克表](#)中选取一些属性，并根据属性相克表查询相克关系，即属性 A 的技能攻击属性 B 的精灵时的加成倍率

- 至少应包括 5 个属性，其两两组合中，“效果拔群”、“效果不佳”和正常效果均应出现至少一次。
- 要求使用技能攻击精灵时，若加成 ≥ 2 输出“效果拔群”， ≤ 0.5 输出“效果不佳”， $=1$ 时不进行输出。

2. 选做部分

1. 增加异常状态：

- 至少包括烧伤、麻痹、中毒、睡眠四个异常状态，并有相应的技能产生这些异常状态（举例：鬼火、电磁波、毒粉、睡眠粉）
- 每只精灵可能因对方的技能而进入异常状态，一只精灵只能同时处于一种异常状态，不可叠加
- 要求在每回合显示双方精灵的异常状态

其中，具体的异常状态如下所示：

状态名称	具体内容
烧伤	攻击减半；每回合结束时损失最大 HP 的 1/16；无法解除
麻痹	速度减半；每回合有 1/4 概率无法使用技能；无法解除
中毒	速度减半；每回合有 1/4 概率无法使用技能；无法解除
睡眠	无法使用技能；持续 1~3 回合（在进入睡眠状态时随机决定回合数）后解除

- 增加种族值：可以对每种精灵设定不同的攻击、防御、速度、最大 HP（、特攻、特防），参考神宝百科具体精灵页面的“种族值”部分
- 丰富能力等级：参见[能力等级](#)
 - 可以考虑区分攻击与特攻、防御与特防（在必做部分中只要求考虑攻击和防御）
 - 可以考虑改变双方攻击、防御、速度、命中率、闪避率的技能
- 增加伤药、恢复药选项：玩家也可以在某回合放弃使用技能，改为使用伤药、恢复药，从而恢复 HP/解除异常状态
- 换精灵：一个玩家可以有最多 6 只精灵；在每回合既可以使用技能，也可以换精灵；若选择换精灵，则在对方的技能发动前完成更换
- AI 进化：对于上述新增内容，你可以设计更高级的 AI
- 增加其他的技能附加效果、更多的技能、特性等杂项。

四、实验指导

1. 设计思路：建议使用**面向对象**的写法，即分别建立精灵类、技能类，对于每个精灵/技能分别实例化精灵类/技能类的对象，设定具体的精灵名称/种族值/属性/...（精灵）或者技能威力/属性/命中率/pp/...（技能）等，对战时每个回合两个精灵分别决定技能（若实现换精灵，则改为用户类决定精灵的技能）

2. 建议对于宝可梦有简单了解的同学考虑本题（~~赛尔号也行，大家应该都玩过~~吧），不然可能需要花较多时间了解游戏机制

3. 参考资料：

- [面向对象基础](#)，一个很短很短的教程，掌握前五部分即可
- [神宝百科](#)。