

“Python 程序设计”课程大作业

任务 1:

分别爬取京东和淘宝，在手机频道中，找出累积销量（所有商家销售同一型号手机的销量之和）最高的 20 款手机。

说明:

- 1) 销量：京东以评论数为准，淘宝以付款人数为准。
- 2) 基础版本：一款手机只看最低配置和最低价格，例如 iPhone 11 只按照 64G 的价格来计算。高级版本：一款手机有多个型号，不同的配置算不同的型号，例如 iPhone 11 64G 算是一种型号，128G 算另一种型号。按照自己的能力选做基础版本或者是高级版本，高级版本有加分。

结果展示:

按照销量，以直方图的形式展示手机型号及其销量，按照销量倒排序。图中的手机型号请尽量简化，能够区分即可。

一个 figure 中有 2 个子图，2 个子图分别展示京东和淘宝中，手机价格直方图。

任务 2:

用你爬取下来的数据制作散点图，横轴为手机价格，纵轴为该价格对应的商家个数。

结果展示:

一个 figure 中有 2 个子图，2 个子图分别展示京东和淘宝中，手机价格的分布情况。

任务 3:

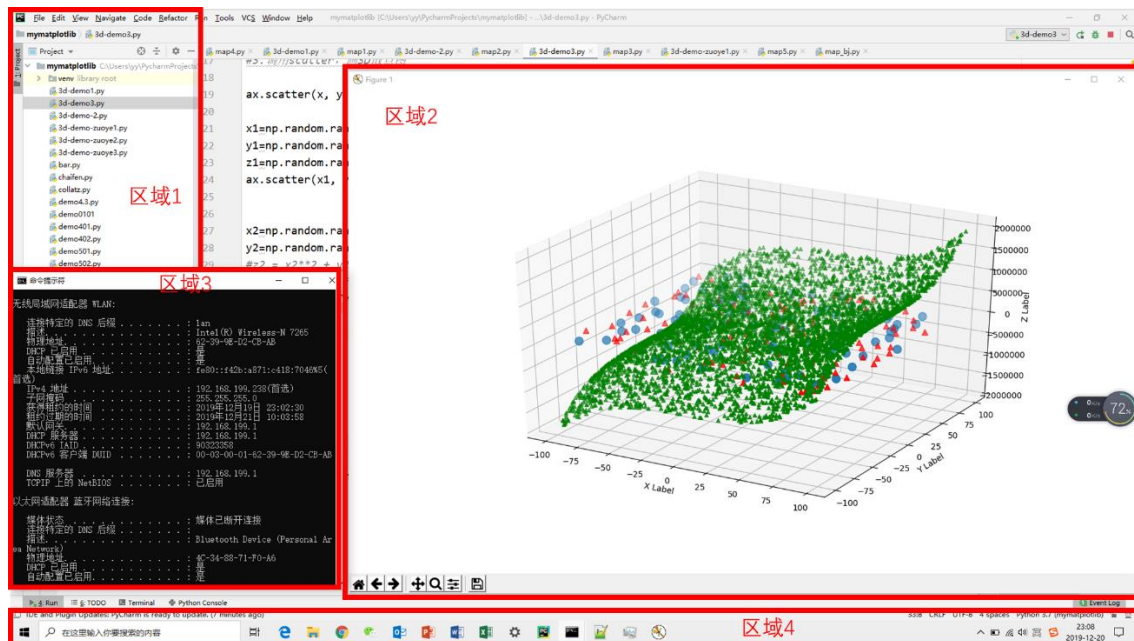
找出在两个平台上都有售卖的 5 款手机（找销量较大的），由于两个平台上都有不同的卖家都销售这些手机，价格也不同，需要将这些卖家销售这款手机的价格，做出箱型图，比较不同平台上的价格情况。

结果展示:

一个 figure 中有 5 个子图，每个子图里面有两个平台上售卖的同一型号手机的 2 个价格箱型图。

提交的 pdf 文档要求:

- 1.pdf 文档以学号命名, 文档内应注明学号和姓名。邮件主题为“python 大作业”。
- 2.请按照以下要求截图:



区域 1: 将工程名称设置为你的学号, 并展示出尽量多的文件。

区域 2: 图形结果展示区

区域 3: 请展示出你的 MAC 地址

区域 4: 请露出这个部分, 不要遮挡。

3.提交核心代码, 例如应包括爬虫的 spider 文件、数据的处理代码、图形的展示代码等。直接粘贴在 pdf 中。

4.如果发现有抄袭现象, 被抄袭和抄袭者均按照 0 分处理。