

Case Study: 个人一卡通收支数据可视化

项目介绍

学校在微信公众号里提供了每个人的一卡通收支数据。



非常方便可以查询自己每个月每天在食堂消费了多少钱：

交易号	时间&地点	金额
264	清真食堂 2018-04-25 12:12:38	-12.0
263	二楼 2018-04-23 17:24:53	-8.2
262	一楼 2018-04-23 08:26:16	-5.5
261	豆腐坊 2018-04-20 17:41:40	-6.0
260	一楼 2018-04-20 17:39:59	-18.0
259	一楼 2018-04-20 17:39:24	-20.0
258	二楼 2018-04-20 11:27:19	-14.3

但是，我们想通过查看这个流水单子，去知道自己的具体消费习惯，却不是那么容易的事。所以这个case study的目的就是将个人收支数据可视化，使其更加直观。这样我们一眼就看出自己每个月在食堂花了多少钱，最钟爱哪个食堂，以及消费的波动情况，总之这个case study就是很有意义的。

动手第一步：下载数据

首先当然是需要数据的，这里需要麻烦一下大家，把2017年9，10，11，12月，2018年3，4月的数据下载下来。请使用微信的电脑客户端，似乎只能用微信自带的浏览器打开流水单子。具体的方法就是将表格复制粘贴到Excel里面（不要包含“交易号”，“时间&地点”，“金额”）。然后保存成Python可以提取的文件格式（.csv utf-8）。这里要注意一点，因为每一页的数据都是最新的在最上面，所以最好是从2018年4月开始，然后2018.3，2017.12到2017.9这个顺序来拷贝，依次往下添加，这样顺序就是正确。

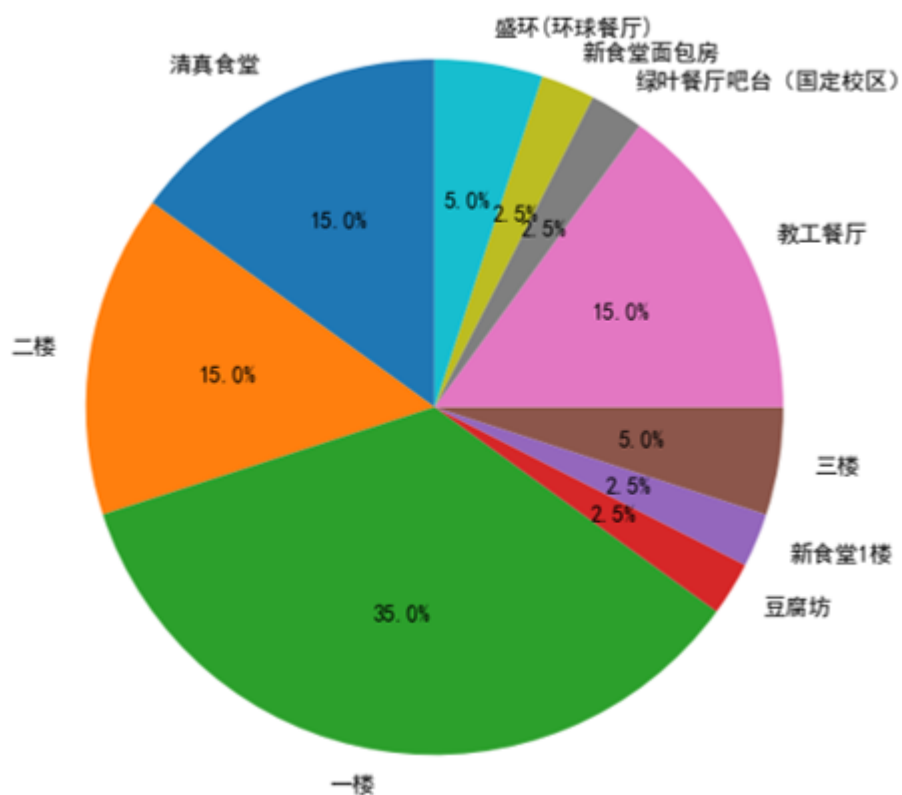
有的同学可能会问，既然都拷到Excel了，为什么不直接用Excel做数据分析呢？当然是因为我们这门课学的是Python，所以要实际操作一下Python咯。

保存成.csv文件后呢，可以用sublime，或者notepad++打开看看，看看是不是整齐的。

动手第二步：整理数据

现在的数据如果被python提取出来是很杂乱的，所以这里我就帮助大家一步，把数据给你们整理好，但是怎么整理的，希望你们好好读一读程序 `case_example.py`。当然我也给出了一个饼状图的例子，也请看清楚，因为后面会用到。

这个饼状图是关于你光顾每个食堂的频率，这里放出我自己四月光顾每个食堂的频率饼状图。

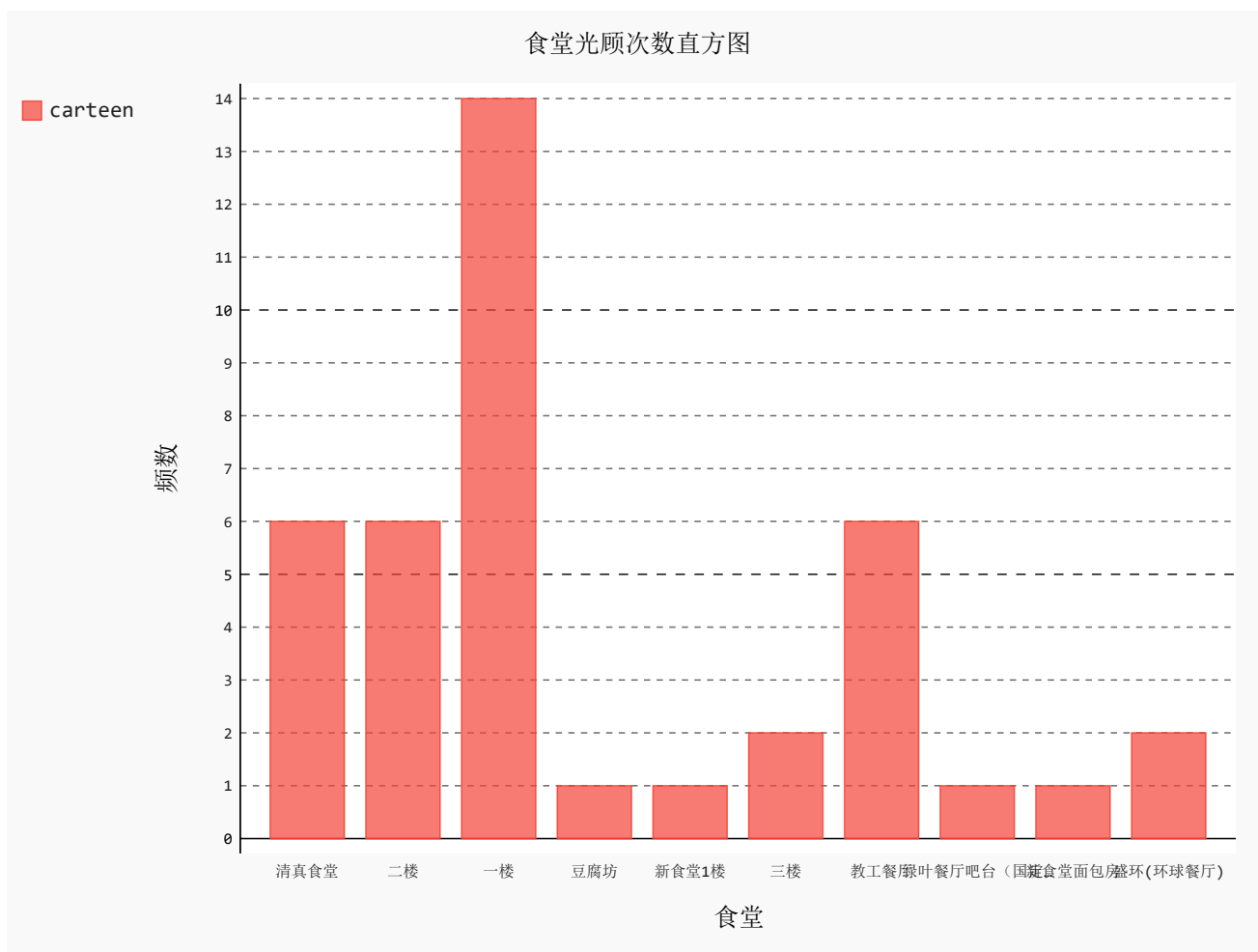


可能做得不够漂亮，但是用来解决我们的问题，应该足够了。因为我所有早餐都是在新食堂一楼吃的，而且充值，宵夜都是在那里比较方便，所以就是新食堂一楼的次数最多了。这么一看就能了解很多信息是不是？

接下来的问题就需要你们自己解决了。

题目一

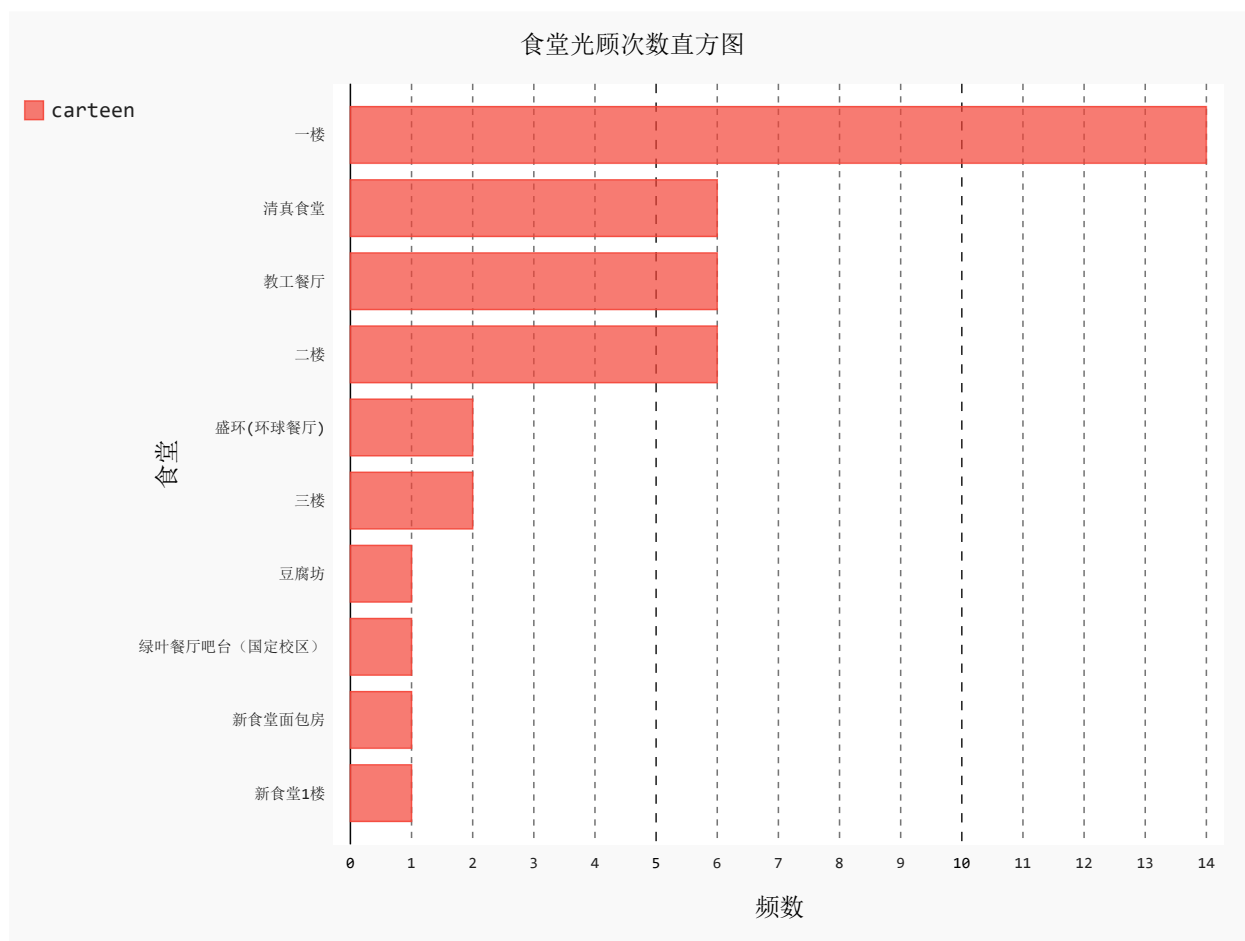
请制作你们六个月里光顾每个食堂的频率直方图 (.svg)。其实只用稍微修改我给出的例子就行了，把饼状图改成直方图，但是要用到你们自己的数据，不是一个月的，是六个月的。下面是我的例子，：



我这个图有两个问题：

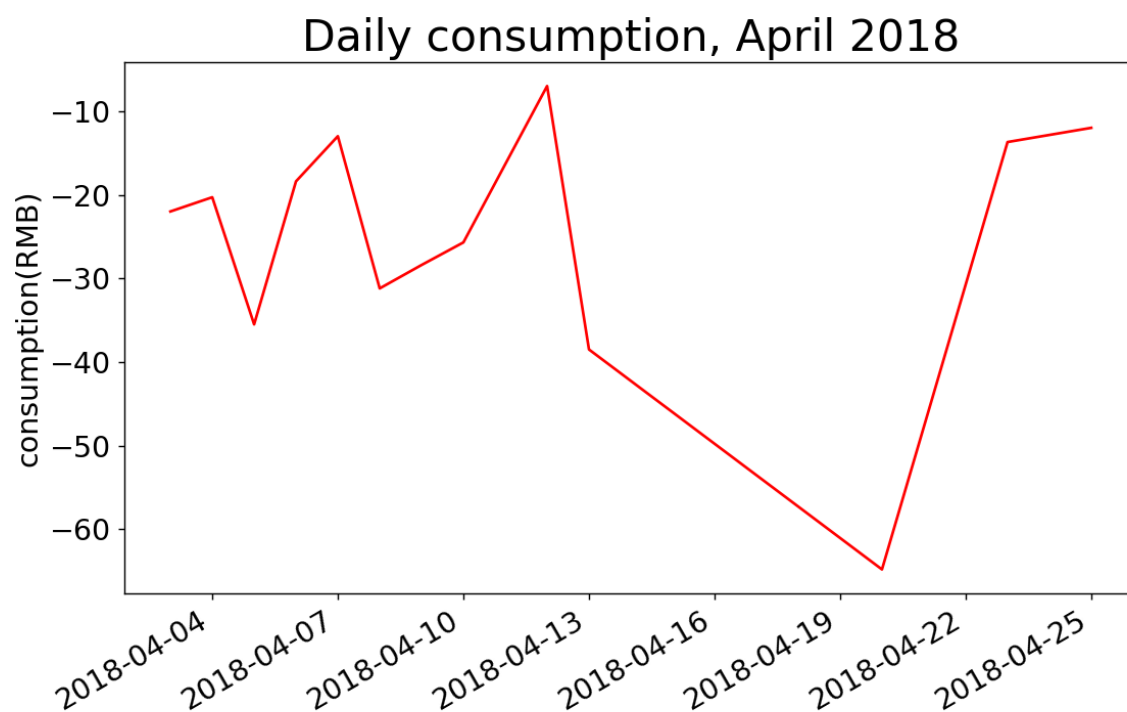
1. 食堂的名字挤到一起去了，大家也会遇到类似的问题，怎么改就交给大家了；
2. 没有从高往低排，如果你觉得从高往低排更好看，就请做成从高往低的样式。

这两个问题也许大家也会遇到，这里就请大家自己想办法解决了，第一个问题建议去查一下Pygal的说明文件(document)。第二个问题就先给频率排个序，怎么排自然是上课讲过的，请各位动动脑筋。当然即使你做不到以上两点，我也是不会扣分的，做到了也不会加分，但是图会变得很好看，就像下面这样：



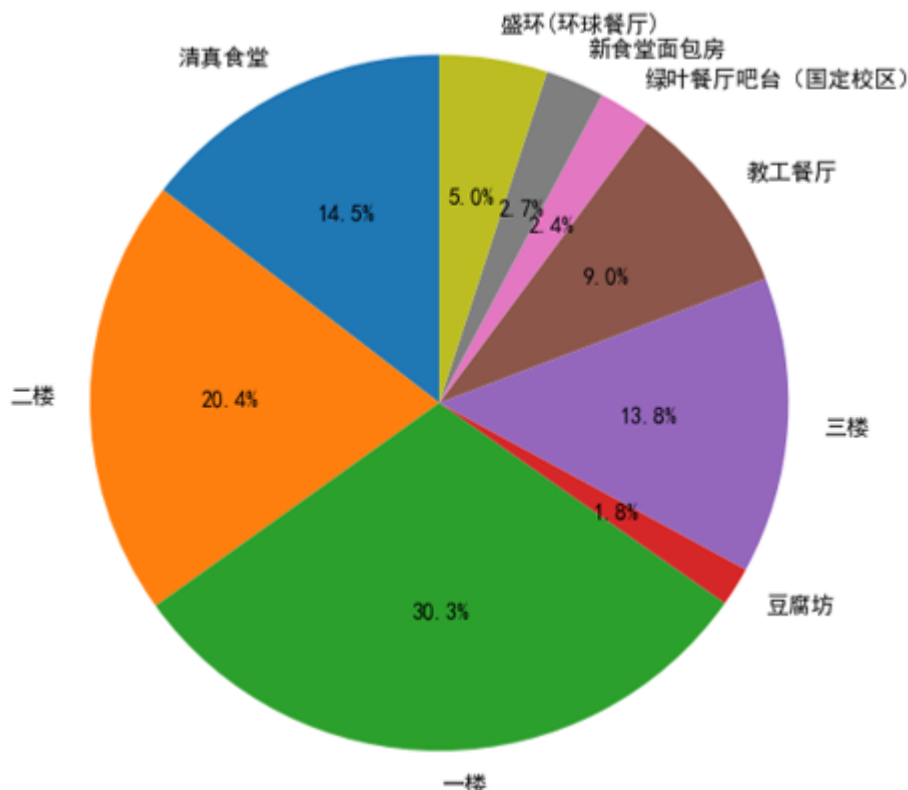
题目二

请制作你们六个月里每天消费额度的折线图。注意两点，1. 请不要把你们充值的额度，当成消费画到图里去了。充值的额度是正的，消费是负的；2. 你们每天有多次消费，应该把一天的所有消费加到一起的。下面是我的例子：



题目三

请制作你在六个月里在每个食堂消费总额度的饼状图。这个题目就是直接运用我给的例子中的饼状图了，但是也要注意，你们必须把在每个食堂的消费加到一起。下面是我的例子：



作业提交

1. 将你们的整理好的数据保存在一个JSON文件里，是整理好的，不是乱乱的哦！
2. 将你们每道题的程序单独作为一个文件，所以一共就是三个Python的文件。
3. 将你们画好的图，放到一个Word文件里。

所以一共就是五个文件: JSON数据文件，三个Python文件，Word文件

结语

当然，这些数据还有很多可分析的空间了，鉴于大家还没有学习概率统计，题目就分析到这里为止，只要你熟练操作Python，就可以以直观的方式，知道你想知道的东西。最后提醒大家一件事：这个case study占期末总分的6%，各位加油！