**练习五——问卷数据分析（10月20日）**

某课题组设计并发布了 《大众点评内容营销对消费者行为影响问卷》，收到有效问卷 219 份，保存在“问卷”文件夹内，每个 txt 文件记录了一个问卷。问卷包含 6 道基础信息（性别、年龄、学历、职业等）和 19 道李克特量表题。数据在GitHub上下载：<https://github.com/MingMingFaery/pythonStudy>

题目一：数据读取：使用 Python 读取并存储所有问卷的数据，求：可支配生活费在3000元以上的人群中，女性所占的比例。

题目二：可视化（1）对样本的年龄分布进行可视化；（2）对不同性别的样本的年龄分布进行可视化；（3）对不同年龄段样本的学历分布进行可视化

题目三：相关性分析：对 19 道李克特量表题结果进行相关性分析，进行可视化并输出相关性最高的两道题

题目四：评价： 19 道李克特量表代表了 19 项评价指标，请使用线性加权的方式对每个样本进行打分，并输出得分最高和最低的样本（使用熵权法对指标赋权）

题目五：分类：有专家根据问卷结果对大众点评用户群体划分了三个类别：A 接受型，B 尝试型，C 厌恶型，目前已知前 200 个样本的类别类型，存储在 Type.csv 文件中，请训练合适的模型，判断其余 19 个样本的类型。根据你的个人观点填写问卷，并使用分类器判断自己属于什么用户类型。