



PROJECT SPECIFICATION

Creating Customer Segments

数据研究

标准	符合规格
问题 1 选取样本	已选取三个数据样本，提出建立表达式并给出合理解释。
问题 2 属性相关性	准确报告被删除属性的预测分数，合理解释被删除属性是否具有相关性。
问题 3 属性分布模式	学生找出具有关联的属性并将其与预测属性相比较，随后深入讨论这些属性的数据分布模式。

数据处理

标准	符合规格
特征缩放	数据和样本的特征缩放已在代码中正确实施。
问题 4 异常检测	学生找出极端的异常值，讨论是否删除这些异常值，并说明删除各数据点的理由。

属性转换

标准	符合规格
问题 5 主要成分分析	准确报告主要成分分析数据的二个维度与四个维度的总方差。将前四个维度合理解释为对消费者支出的表达。
降维	对二维缩放数据及样本数据的主要成分分析已在代码中正确实施。

聚类

标准	符合规格
问题 6 聚类算法	高斯混合模型和K-均值算法已进行详细比较。学生选择的算法符合算法和数据的特点。

标准	符合规格
问题 7 创建集群	准确报告多个轮廓分数，根据报告的最佳分数选择最佳集群数量。已给出的集群可视化将根据已选的聚类算法生成最佳的集群数量。
问题 8 数据恢复	根据数据集的统计描述提出每个客户细分所代表的类型。对集群中心的逆变换和反比例级联已在代码中正确实施。
问题 9 样本预测	客户细分正确识别样本数据点，讨论各样本数据点的预测集群。

结论

标准	符合规格
问题 10 A/B 测试	提出了某些功能改进方法，可以改进从 A/B 测试获取结果的功能。
问题 11 预测额外属性	学生讨论了聚类数据如何可以通过监督学习预测新的属性。
问题 12 比较客户数据	客户细分与客户通道数据进行对比，对通道数据识别客户细分的问题进行讨论，包括该表达是否符合早期结果。

Suggestions to Make Your Project Stand Out!

你可以在[这里查看](#)此评审标准的英文版本。