

问题：销售数据分析

背景：模拟销售分析

假设你现在是一家连锁商场某分店新上任的销售经理，需要了解该分店过去一段时间的销售业绩情况，并根据数据分析结果制定相应的销售计划。

附件说明：

1. 附件为5个文件，其中1-4为数据文件，5为对应的数据字典及字段说明。
2. 数据会存在一定内容缺失，请描述缺失数据的处理操作，合理即可。
3. 附件1. 该分店注册的会员信息表。
4. 附件2. 该分店的总销售记录。（本店会员，非本店会员，非会员的销售记录总和）
5. 附件3. 该分店会员的销售记录。（本店会员，非本店会员的销售记录总和）
6. 附件4. 商品信息表。
7. 附件5. 数据字典。

要求：

1. 注意审题。
2. 请通过文字分析4张表的字段信息，指出可以得到哪些直接，间接的信息。
3. 统计该商场的总销售额，以及按年度，月度统计销售额。分析销售额的变化是否有周期性。
4. 该店有效会员的性别，年龄，及注册时间情况分析。请自行理解“有效会员”的描述，并给出说明。
5. 该店有效会员的消费金额统计。并绘制消费金额分布图。
6. 通过分析有效会员的性别，年龄，入会时间，消费总额，消费笔数等信息，尝试分析哪些会员才是有价值的会员。给出分析与结果。
7. 如有余力，可以尝试完成或扩展以下要求。
8. 按品牌，按类别统计销售额（总体，年度，季度，月度等）及排序。也可选择某类单独分析。
9. 分析哪些品牌或类别是受性别，年龄影响较大？
10. 其他。

补充：

1. 使用相对路径载入数据文件。即数据文件与.py文件在同一目录下。
2. 请提供完整的可运行代码.py，及对应的书面报告。代码请加注释说明。附件无需提交。
3. 尽可能使用简洁高效的代码。
4. 注意编码，变量名，函数名，类名等命名规范。
5. 开发语言限制为Python或者R。
6. **欢迎在完成要求后拓展分析。**
7. 本题目是一道综合性较强的数据分析题。包含了基础数据的清洗，多表字段关联分析，实际应用分析，基本的数据处理变形操作，绘图操作，甚至一些统计方法等多方面的内容。除了基础统计外，所有的分析结果没有标准答案，考察的是被考察者根据题干要求与数据信息做出的合理假设与分析过程。希望你在完成或尝试完成本题的过程中，让你有所收获。
8. 本题考察的是：代码书写习惯，语法掌握情况，逻辑分析水平，建模能力。