

预测宣传册需求

请完成每个部分。准备好后，将你的文件另存为 PDF 文档并从课堂上提交。

第 1 步：理解业务和数据

解释下需要作出的关键决策。（限 500 字以内）

关键决策：

请回答以下问题

1. 需要作出什么样的决策？

答：针对 250 名新客户，预期利润是多少，是否达到要求（利润超过一万美元），如果达到要求，则寄送产品目录册。

2. 作出这些决策需要获取哪些数据？

答：Avg Sale Amount, Customer Segment, Avg Num Products Purchased

第 2 步：分析、建模和验证

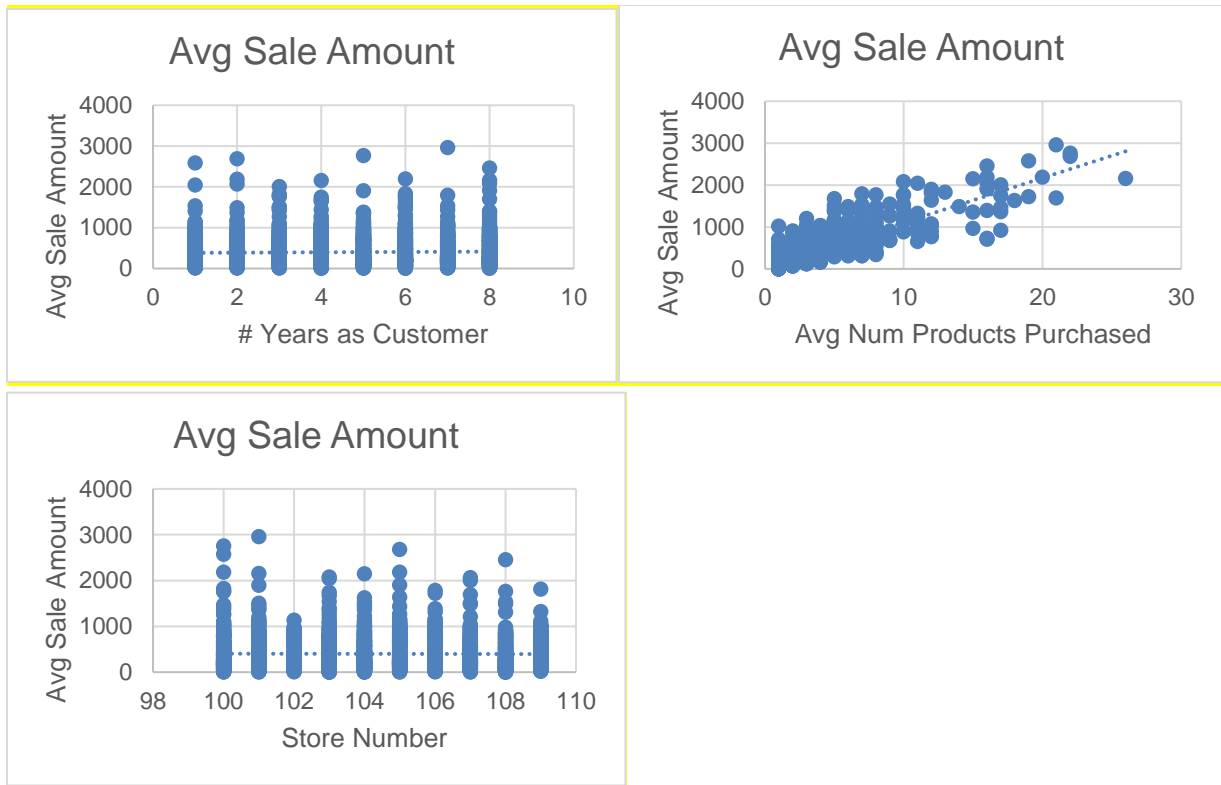
描述下你是如何设置线性回归模型的，使用了哪些变量，原因是什么，以及模型的结果。建议提供可视化图表（限 500 字以内）。

重要事项：使用 **p1-customers.xlsx** 训练你的线性模型。

至少回答以下问题：

1. 你是如何在你的模型中选择预测变量的？原因是什么？你必须解释你选择的连续预测变量与目标变量有线性关系。请参阅“线性回归-Excel 的多元线性回归”来探索你的数据，并使用散点图寻找线性关系。你必须在答案中包含散点图。

答：预测 Avg Sale Amount，分别选取了 Avg Num Products Purchased, # Years as Customer, Store Number 等变量与 Avg Sale Amount 分别绘制散点图：



可见，Avg Sale Amount 与 Avg Num Products Purchased 有线性关系。

- 解释为何你认为你的线性模型是很好的模型。必须使用你的回归模型产生的统计学结果证明你的推理过程。对于你所选择的每个变量，请使用你的模型产生的 p 值和 R 平方值证明每个变量为何与你的模型很好地拟合。

回归统计	
Multiple R	0.91481
R Square	0.836878
Adjusted R Square	0.836602
标准误差	137.4832
观测值	2375

$R^2=0.836878$, >0.7 , Adjusted $R^2=0.836602$, >0.7 线性模型解释性较强。

	Coefficients	标准误差	t Stat	P-value
Intercept	303.4634713	10.57571	28.69436972	1.1227E-155
Avg Num Products Purchased	66.97620492	1.51504	44.20753848	0
Store Mailing List	-245.4177445	9.767776	-25.12524388	1.0503E-123
Loyalty Club and Credit Card	281.8387649	11.90986	23.66432739	2.5804E-111
Loyalty Club Only	-149.3557219	8.972755	-16.64547014	6.34584E-59

P 值都 <0.05 ，其与目标变量之间的关系被认为具有统计学意义。

3. 根据提供的数据，最佳线性回归方程是什么？每个系数小数点后最多保留两位（例如 1.28）

$$\text{Avg Sale Amount} = 303.46 + 66.98 * \text{Avg Num Products Purchased} - 245.42 \text{ (if type: Store Mailing List)} + 281.84 \text{ (if type: Loyalty Club and Credit Card)} - 149.36 \text{ (if type: Loyalty Club Only)}$$

第 3 步：演示/可视化:

根据你的模型结果给出建议。（限 500 字以内）

至少回答以下问题:

1. 你的建议是什么？公司应该向这 250 个客户发送宣传册吗？

答：250 个新客户的预期利润是 21987.96 美元，大于 10000 美元，建议向这 250 个客户发送宣传册。

2. 你是如何得出你的建议的？（请解释你的推理流程，以便审核人员能够根据你的流程向你提供反馈）

答：

- 根据线性回归模型计算 250 个新客户每个新客户的预期 Score
- 计算每个新客户的预期利润（假设寄送宣传册）：预期 Score*50%-6.5
- 计算 250 个新客户预期利润总数
- 250 个新客户预期利润总数>10000，建议寄送宣传册

3. 新的宣传册带来的利润预计是多少？（假设向这 250 个客户发送了宣传册）

答：250 个新客户的预期利润是 21987.96 美元。

提交之前

请根据此处的[审核标准](#)中列出的项目要求检查你的答案。审核人员将根据该审核标准对项目打分。