

学习效率数据分析

Razonado sobre formato de datos,持续 17 秒 从你提供的截图来看,你的表格包含了以下信息:

- 日期、开始/结束时间
- (尚未明确的) 学习总时长
- 任务数、任务类型、任务评分
- 学习内容或质量描述
- 休息情况、睡眠情况
- 情绪及身体状况评分(看起来既有文字描述也有数字)
- 使用策略及工具

总体而言,这些字段对研究学习效率和可持续性来说非常有帮助,已经包含了时间、内容、主观状态等多个维度。为了后续的数据分析更加便利,建议你可以做以下几点优化:

1. 将「开始/结束时间」转换为可计算的「学习总时长」

- 单独存储开始/结束时间:目前你已经有了「开始学习时间」「结束时间」两个字段,但还没有清晰的数值型「学习总时长」。
- **计算方式**:可以在 Excel 中增加一列「学习总时长(小时)」或「学习总时长(分钟)」:

学习总时长 = 结束时间 - 开始时间

如果结束时间小干开始时间(如跨天情况),则需要特殊处理。

• **注意格式**:要保证「开始/结束时间」在 Excel 中是可以被识别为时间格式,然后再做差值。

2. 将「休息情况」「睡眠情况」标准化

目前你用的是文字描述(例如「中午休息但未睡着」「8个小时睡眠」「睡眠质量较差」等),后期做量化分析 时会比较麻烦。可以考虑:

- 1. **休息时长**:例如在「中午休息」列中填入具体时长(分钟或小时)。如果没有午休则填 0,或者填具体时间段。
- 2. **睡眠时长**:用一个数值列来记录睡眠时长(单位小时),再用另一列来记录主观的睡眠质量评分(15 或 110)。
- 3. **休息/睡眠描述**:如果你还想记录文字感受,可以在备注列中写更详细的文字描述。

这样既能有量化指标,也能保留原始描述供参考。

3. 将「情绪及身体状况」拆分为「情绪评分」和「身体状况评分」或单一量化

目前看到的内容类似「低落,疲劳3」「心态平和,稍微紧张」。这会在数据分析时比较难直接做回归或相关分析。

● 你可以使用一个**单一量表**(比如 15 或 110)来综合衡量情绪和身体状态,也可以拆分成两个维度:

- 情绪评分:1(非常糟糕)~5(非常好)
- 身体疲劳度:1(精力充沛)~5(极度疲劳)
- 如果你想要更多维度,也可以扩展为「焦虑」「疲劳」「愉悦」等几个分量表,但要考虑到记录的便捷 性。

4. 为「任务质量/正确率」和「任务评分」建立可对比的量表

- 「质量 or 正确率」目前是文字描述,如「看完三节网课」「主要时间用来熟悉工具」等。
 - 可以尝试用一个分数或完成度指标来量化(如「网课完成率 0100%」或者「理解度 15」)。
- 「任务评分」有时是数字(如 0、3、1、3),但并不清楚评分标准是什么。
 - 如果想要在后续进行相关或回归分析,需要明确评分的含义(例如 0 代表没完成,5 代表超额完成? 还是表示质量?)。
 - 最好在表格里备注或在你的研究设计文档中说明评分标准。

5. 对「任务类型」「使用策略及工具」进行分类编码

- 「任务类型」可以先定义几个类别(如:阅读、写作、编程、上课、预习、复习等),然后用**分类编码**的 方式记录在 Excel 中(如数字编码或者独热编码 One-hot encoding)。
- 「使用策略及工具」同理,可以先把常见的工具和策略(如「番茄闹钟法」「无方法」「思维导图」等) 列成一个表,分配编码值,这样后续做机器学习时比较容易处理。

6. 补充或优化的可选字段

- **学习目标或当日重点**:有时记录当日的具体目标可以帮助在分析时区分「完成度」。
- **专注时长**:如果你用番茄钟等工具,可以记录每天专注时间(Pomodoro 个数或专注时长)。
- **干扰次数**:比如记录你当天在学习中被打断或走神的次数,也会影响学习效率。

7. 确保记录的持续性和一致性

最后,要想进行有效的相关分析或建模,需要有**足够多的记录天数**和**一致的记录方式**。建议:

- 尽量在每天的固定时间更新数据。
- 尽量使用相同的量表或评分方式,避免过于随意的描述,保证数据的可比性。

总结

你的表格格式整体上已经具备可分析的基础,只是在细节上还需要做一些**标准化和量化**的工作,以便后续做回归分析、相关分析或机器学习模型时更方便、更准确。

如果你能根据以上建议对表格做适当调整并持续记录一段时间,那么我们就可以更容易地使用如 Python(pandas、scikit-learn)或 Excel 统计功能来进行回归、聚类或其他类型的分析,最终尝试找到影响学习效率和可持续性的主要因素,并给出一定的预测公式。

有任何问题可以随时告诉我,我会继续为你提供更具体的指导和示例。