

作业: 方差分析(One-Way ANOVA)

介绍

一个在线零售商希望提高员工的工作效率，同时改善他们的工作体验。目前，零售商订单管理中心的员工在工作时没有得到任何形式的娱乐，如背景音乐、电视等。零售商想知道提供一些员工要求的播放背景音乐是否会提高生产力，如果能提高，具体能提高多少。

因此，一名研究人员随机抽取了 150 名员工。这 150 名参与者被随机分为三组，每组 50 名参与者：（a）一个“对照组（control group）”不听音乐；（b）一个“治疗组（treatment group）”，他们听音乐，但不能选择听什么；和（c）第二个“治疗组”（treatment group），他们不仅可以听音乐，还可以自主选择听什么音乐。

实验持续一个月。在实验结束时，三组的“生产力”是根据“每小时处理的平均包裹数量”来衡量。因此，因变量是“生产力”：由一个月实验期间每小时处理的平均包装数量衡量，而解释变量是“分组”：有三个相互独立的组，“no_music”（对照组）、“music_no_choice”（治疗组 A）和“music_choice”（治疗组 B）。研究人员假设，适当的娱乐放松会提高生产力，处于可选择音乐工作状态下的员工（治疗组 B）的生产力水平最高，其次是不可选择音乐工作状态下（治疗组 A），最后是无音乐组（对照组）。

请撰写报告，回答下面这些问题。

- 1) 选择的数据分析方法及选择理由
 - 如何分析数据，以及阐述为什么选择这种分析方法
- 2) 基础分析
 - 数据筛选、数据分布分析
- 3) 主要分析
 - 采用统计分析方法（如 ANOVA）开展主要数据分析工作，形成结论，提供必要的表格和图表
- 4) 对下面研究问题的解读
 - 通过上面的分析结果来回答下面的研究问题

研究问题

- 1) 这三组之间的生产力是否存在统计学上的显著差异？
- 2) 两两来看，组之间的生产力是否有区别？哪个组具有最高的生产力水平？

写作格式

报告内容至少包括以下几个部分

[标题]

[研究问题和假设]

[选择的数据分析方法和理由]

[基础分析]

[主要分析]

[分析结果解读]

数据文件

music_data.csv