

我要如何完成该项目？

该项目与数据分析入门课程相关，但根据你的知识背景，你可能无需学习全部课程即可完成该项目。

介绍

在此项目中，你需要进行数据分析，并创建一个文档分享你的发现。你应该首先了解数据集，想想它可以用来回答哪些问题。然后，你应该使用 **Pandas** 和 **NumPy** 回答你最感兴趣的问题，并编写一份报告来分享你的结论。你不需要使用推论统计或机器学习知识来完成此项目，但是你需要在报告中声明，你的结论是暂时的，可能需要进一步改进。这个项目是开放性的，没有标准答案。

第一步 - 选择数据集

点击[此链接](#) 打开一个文档，其中包含你可以为此项目调查的数据集的链接和信息。你**必须**选择其中一个数据集来完成项目。 若数据无法下载，可以通过以下链接进行下载：

1. [TMDb 电影数据](#)
2. [未前往就诊的挂号预约](#)
3. [Gapminder World](#)
4. [足球数据库](#)
5. [FBI 枪支数据](#)

第二步 - 组织

最终，你需要提交一份报告（并与你的朋友、家人或同事分享）。在正式开始之前，让我们进行组织整理。我们建议你新建一个文件夹，它最早将包含：

- 用来传达你的发现的**报告文档**
- 你用来分析数据的任何 **Python 代码**
- 你使用的**数据集**（你不需要提交它）

你可以使用 **Jupyter Notebook**，这样，你可以在同一份文档中提交编写的代码和报告你的发现。不然你可能只能分别提交报告和代码了。如果你需要一个**笔记本模板**来帮助组织你的调查，可以在页面底部找到资源链接或可以点击[此处](#)。

第三步 - 分析数据

思考你可以用数据集回答哪些问题，然后开始着手回答这些问题。你可以在[数据集选项](#)中找到一些问题，以帮助你入手。试着提出那些调查多个变量之间关系的问题。你应该在你的调查里分析至少一个因变量和三个自变量。在恰当的地方使用 **NumPy** 和 **Pandas**。

第四步 - 分享发现

分析完数据以后，编写一份报告，分享你觉得最有意思的发现。如果你使用 **Jupyter Notebook**，则与你用于执行分析的代码一起分享发现。确保你的报告文本包含在 **Markdown** 单元中，以清楚地将你的注释和发现与代码分开。当然你也可以使用其他工具和软件来制作最终报告，但请确保你能以 **HTML** 或 **PDF** 文件的形式提交报告，以便轻松打开。

第五步 - 检查

使用[项目评估准则](#)来检查你的项目。如果你的提交已达到所有要求，则可以提交项目。如果你发现还有改进的空间，则继续改进吧！