

项目：可视化电影数据

你需要完成每一个部分，当你准备好了，将文件另存为 PDF 文档格式，在教室里提交。

第一步：清理数据和选择变量

- 清理缺失的信息，选择在可视化中需要进一步研究的最重要的变量。
- 列出你在可视化中会进一步研究的变量。你要研究的变量数量应该不超过 8 个。
Production_companies,release_year,genres,keywords,popularity,budget,revenue。
- 可以使用 Python 中的 Pandas 库来清洗数据。

第二步：Tableau 可视化

- 请确保你遵循了这些[提示](#)，并在你的项目中包含了 Tableau 的仪表板（Dashboard）、故事（Stories）模块，以及你的项目评审师会特别检查的可视化类型（小图、散点图、条形图等）。需要特别注意的是，针对每一个问题（Q1-Q4），你都需要创建一个仪表板。另外，你也需要为一个你所选择的问题创建故事模块。

重要提示: 请把你的工作簿上传到 **Tableau Public** 上，并允许你的项目评审师查看你的工作簿。请注意单单把文件简单保存为后缀名为“.twbx”不能让所有项目评审师查看文件。[如何做到能够允许所有项目评审师查看文件的指示在这里。](#)

第三步：问题

- 回答下列问题，引用你在线可视化结果去支持你的答案：
 - 问题 1:** 电影类型是如何随着时代变化而变化的？
 - 一是从数量方面分析：随着时代变化，大部分电影类型的数量不断增加，其中 Drama 类型电影变化最大，2014 年达到巅峰 117 部，2009 年 Comedy 类型电影达到最高 103 部。
 - 二是从点击量方面分析：Action 类型电影一直被人们高度关注，点击量越来越高，其次是 Drama 类型电影，第三是 comedy 类型电影。
 - 三是从收入方面分析：Action 类型电影表现非常强势，收入逐年提高。
 - 链接：https://public.tableau.com/shared/PP86Y34RB?:display_count=yes
 - 问题 2:** 环球影业和派拉蒙影业的电影之前数据指标有什么区别？
 - 一、1990 以来，环球影业电影数量一直超过派拉蒙影业；
 - 二、环球影业和派拉蒙影业平均收入基本相同，但环球影业总收入超过派拉蒙影业；
 - 三、环球影业和派拉蒙影业平均点击量相同，但是环球影业总点击量超过派拉蒙影业。

链接:

https://public.tableau.com/views/question_two/q2?:embed=y&:display_count=yes

问题 3: 和非小说改编的电影相比, 基于小说改编的电影表现得怎么样?

首先非小说改编电影在数量上占据了绝对的优势, 小说改编电影数量表现的差强人意;

其次小说改编电影预算和收入波动比较大, 而非小说改编电影的比较稳定, 整体上小说改编电影平均预算和收入高于非小说改编;

最后小说改编的平均点击量波动非常大, 非小说改编平均点击量相对稳定, 但小于小说改编。

链接:

https://public.tableau.com/views/question_three/1_1?:embed=y&:display_count=yes

- 你提出的另外问题是什么? 答案是什么? 你是怎么想出这个问题的?

问题 4: 电影行业前景如何?

一是电影行业整体上盈利大于亏损, 其中盈利电影占到 **76.5%**, 亏损电影占到 **23.5%**。

二是随着年代发展, 电影行业盈亏都呈现了上升的趋势, 但时盈利趋势相对亏损更为陡峭, 尤其是近 10 年内, 盈利为指数上升的态势, 也就是说目前电影行业发展非常好。

提出此次问题的原因: 当前社会除了房地产行业外, 各个行业前景不是非常乐观, 因此提出了这类问题。

链接:

https://public.tableau.com/views/question_four/q4?:embed=y&:display_count=yes&publish=yes