

项目：可视化电影数据

第一步：清理数据和选择变量

在 excel 中修改错位的数据，利用筛选功能清洗数据，对 genres 变量分列，然后导入 tableau，利用 tableau 数据清洗器再次清理数据，并对 genres 变量透视，以便于后续分析。最后，确定在可视化中会进一步研究的变量有 genres、budget_adj、revenue_adj、vote_average、release_date、key words、production company、director。

第二步：Tableau 可视化

https://public.tableau.com/views/BA3_movies/Q1?:embed=y&:display_count=yes&publish=yes

第三步：问题

- **问题 1：**电影类型是如何随着时代变化而变化的？

1990 年之后，电影类型越来越丰富，但近几年有集中于几个热门类型的趋势。我们可以查看不同类型电影的数量和票房随时间的变化。可以看出，近些年动作、犯罪等类型是拍摄热门选项，且票房表现较好。

- **问题 2：**环球影业和派拉蒙影业的电影之间数据指标有什么区别？

两个电影公司在电影类型选择上各有偏好，而且并非完全基于利润考虑。整体而言，派拉蒙影业对电影的预算投资高于环球影业，票房收入也更好。

- **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？

整体而言，小说改编的电影预算较高，但票房也较高，净盈利额高。同时，小说改编的电影评分集中，好评较多。电影评分和票房并无明显关系，不过票房高的电影一般评分也较高。

- **问题 4：**不同导演拍摄的电影水平和票房有何差异？

我们已经了解了电影风格的选择，并考虑改编小说，还向优秀电影公司学习，下一步我们应该了解导演的情况，以确定选择哪些导演来制作我们的电影。通过筛选出单人执导的电影平均净盈利最高的前 10 位导演，可以看出他们的成功不仅来自风格的选择，主要还是由于电影质量较好。根据电影质量，我们可以选择 J. J. Abrams 和 David Yates 执导我们的电影。