# **项目：可视化电影数据**

## 第一步：清理数据和选择变量

|  |  |  |  |
| --- | --- | --- | --- |
| 数据项 | 数据名称 | 数据来源 | （进一步）解释 |
| 1 | popularity | Movie data | 作为一个度量值在Tableau中研究 |
| 2 | keywords | Movie data | 作为一个维度值在Tableau中研究 |
| 3 | runtime | Movie data | 作为一个度量值在Tableau中研究 |
| 4 | genres | Movie data | 作为一个维度值在Tableau中研究 |
| 5 | Production company | Movie data | 作为一个维度值在Tableau中研究 |
| 6 | Release date | Movie data | 作为一个维度值在Tableau中研究 |
| 7 | Vote average | Movie data | 作为一个度量值在Tableau中研究 |
| 8 | Budget\_adj | Movie data | 作为一个度量值在Tableau中研究 |
| 9 | Revenue\_adj | Movie data | 作为一个度量值在Tableau中研究 |

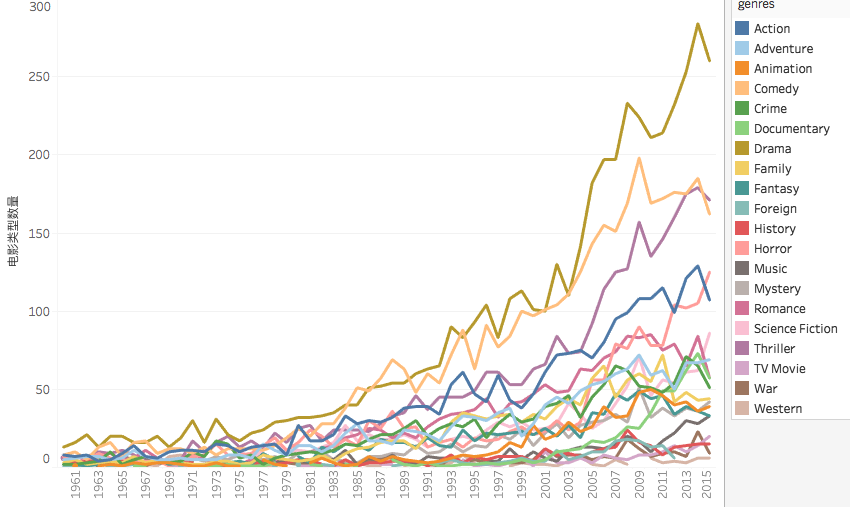
## 第二步：Tableau 可视化

Tableau Public地址：

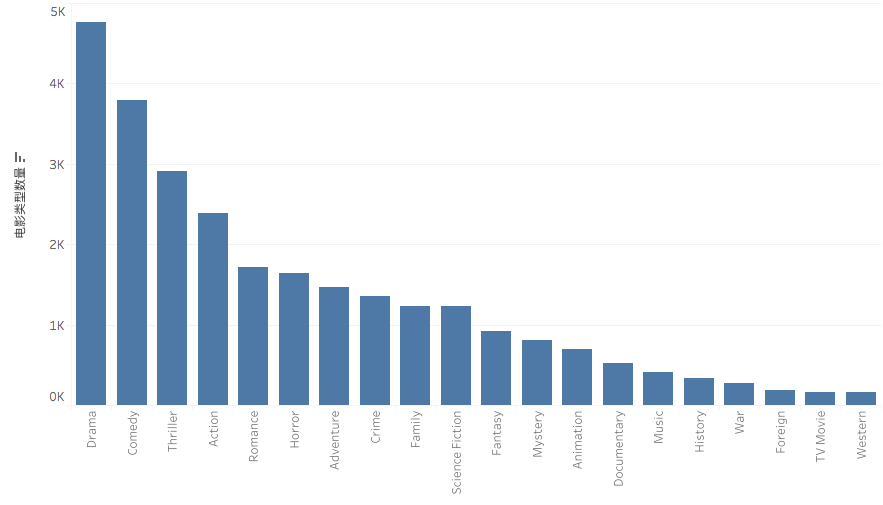
https://public.tableau.com/profile/wenya.chu#!/vizhome/Project3---movie/Q3?publish=yes

## 第三步：问题

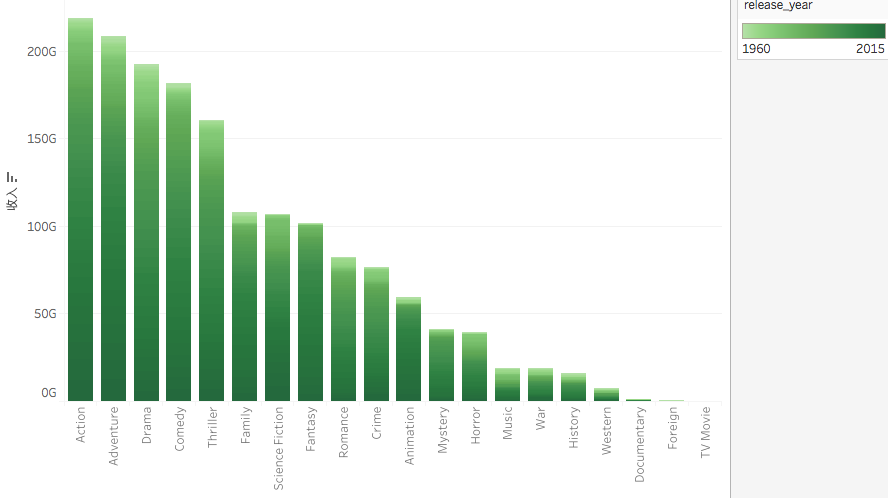
**问题1：**电影类型是如何随着时代变化而变化的?



从图中可以看出，首先，所有类型的电影数量都呈显著增加趋势，这其中最为明显的是Drama，Comedy和Thriller三种

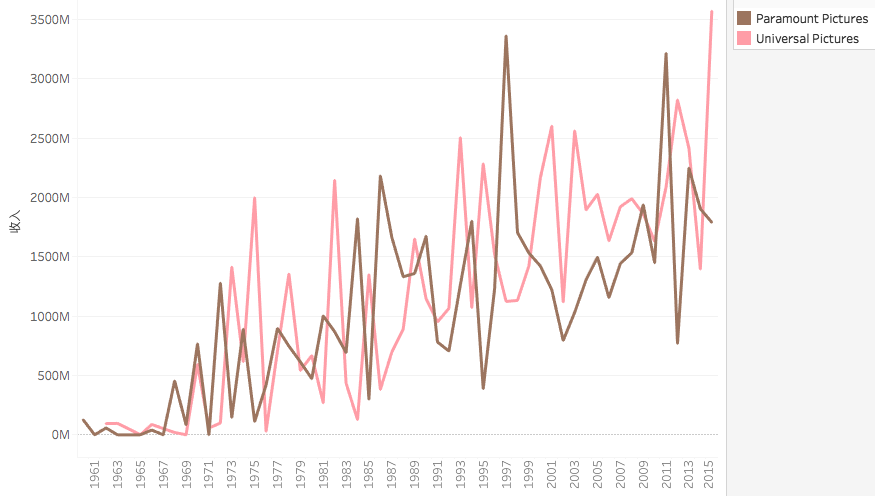


这个直方图也印证了折线图所显现的趋势，Drama，Comedy和Thriller三种类型的电影是最受欢迎，拍摄数量最多的

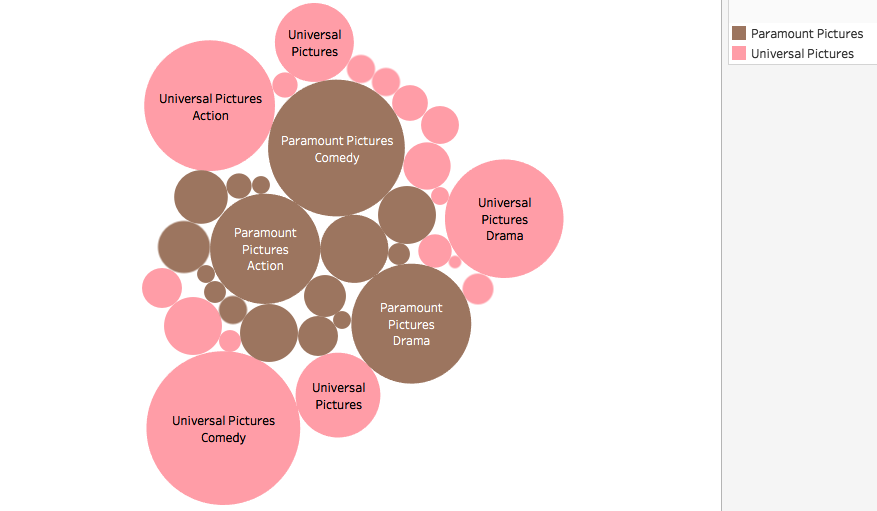


而这张关于各种类型电影的票房收入的图则说明了一个有趣的现象，票房最高的电影类型并不是数量最多的Drama，Comedy和Thriller，而依次是Action, Adventure和Drama

**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?

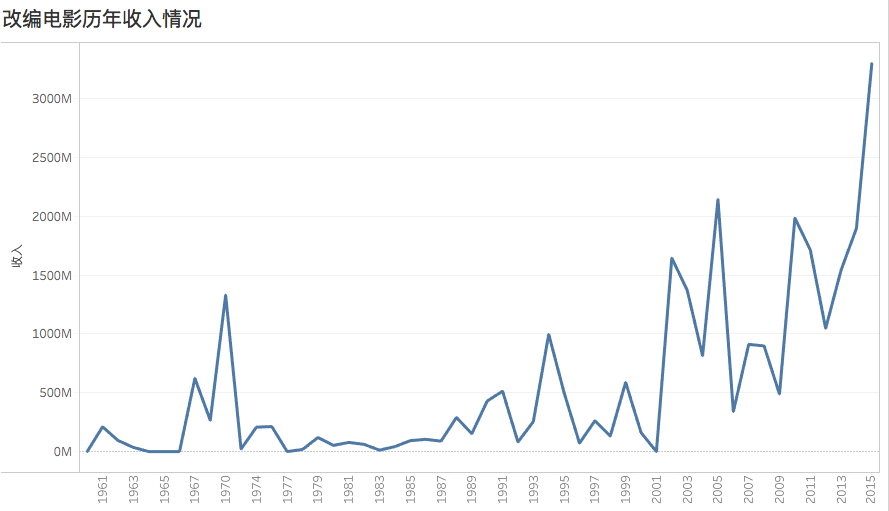


上图是两家公司历年收入情况对比，整体上来说， Paramount Pictures的收入情况波动较大，最高值和最低值都比较明显，而Universal Pictures的收入情况整体呈现逐年上升趋势

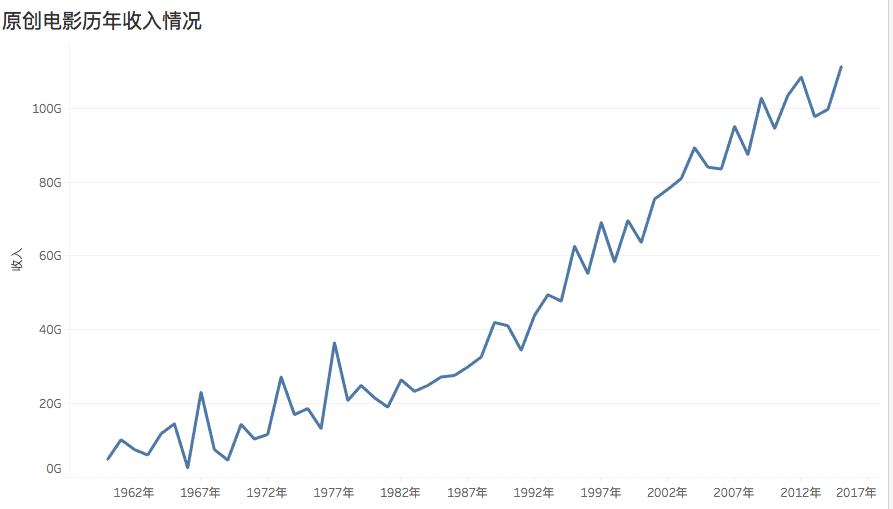


从上图比较两家公司历年的电影类型统计情况，两家公司前三位的类型片都是Comedy, Action, Drama, 但是Universal Pictures的统计数量要多于Paramount Pictures

**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?

****

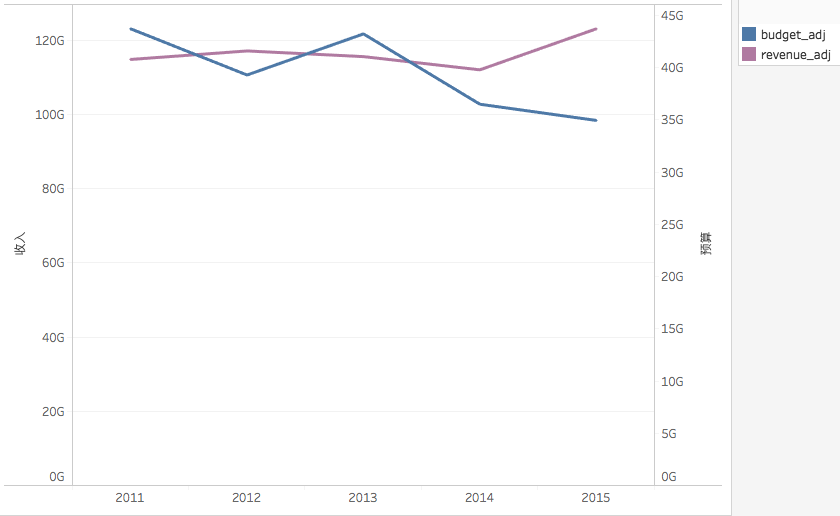
从上图可以看出，改编电影的收入在1970年达到一个小高峰后一直保持低落状态，直到2000年后开始爆发，一直保持很大的上升趋势，并在2015年突破30亿的收入水平。



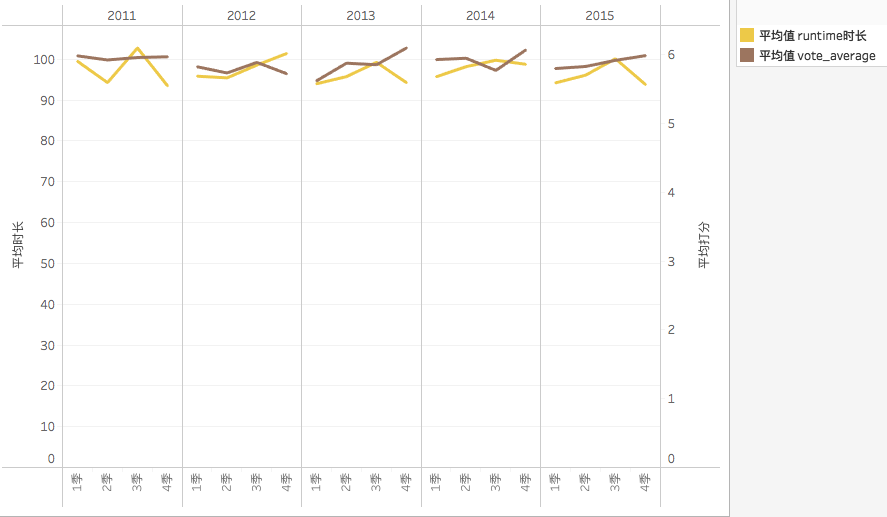
从上图可知，原创电影的整体收入呈现逐渐上升趋势，上下浮动虽有但是不大，并且由于原创电影的类型和数量要远大于改编电影，故收入也不在一个数量级别，原创电影在2017年已经突破100G大关，远超改编电影。

**问题4：**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？

我的问题是01~05年电影市场的情况变化是怎样的？因为我对最近几年的电影市场现状感到好奇，所以想要探索这个问题。

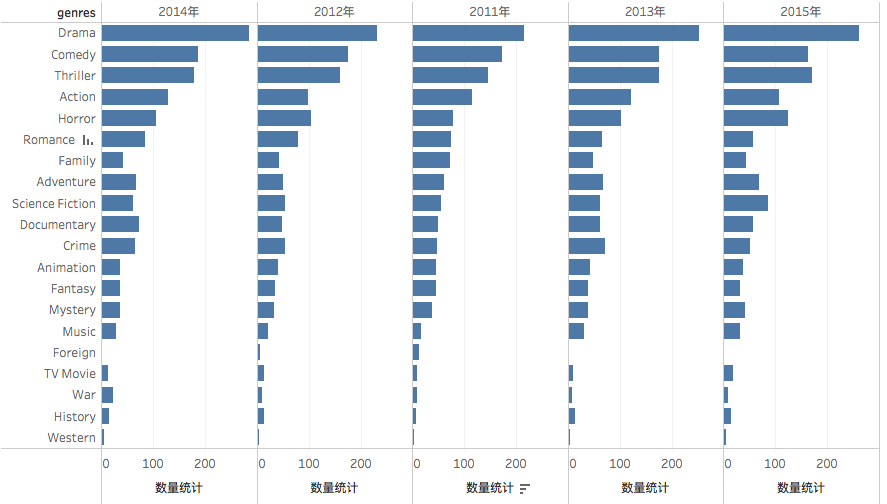


首先我研究了五年间电影市场的预算和收入情况对比，可以看出大体上来说，预算虽然在12，13年有较大波动，但大致还是呈下降趋势，票房收入逐年上升，已经突破120G



接下来我试图探索这五年间的电影时长有何变化，以及观众对于电影的打分情况如何，同时也想找出两者之间是否会有关系？从图中可以看出，我将每年都分为四个季度来处理，便于更清晰细致探索数据。01年电影时长波动较为明显，但评分一直很稳定；02年电影时长一直上升，但评分反而下降；13年的电影时长从第三季度开始下跌，与之相反的是评分反而上升不少；14年第三季度时长增加，评分降低，第四季度时长减少，评分升高；15年时长呈抛物线，但评分一直稳步上升。

总结一下，时长和评分之间似乎存在某种关系，并不是电影时长越长评分就会越高，相反，根据这五年的数据，电影时长越短，评分反而会越高。



这幅图中主要讨论的是五年间电影类型的变化，可以看出，总体上来说，最受欢迎的电影类型一直是Drama, Comedy, Thriller，其他类型的电影虽然数量上有波动，但变化都不大。