



# 影响电影票房相关因素分析报告

## **Analysis of the factors affecting the box office of films**

报告人：赵哲冕 17420182200850

李文豪 17420182200748

邵新哲 19020182203621

二〇二一年六月十日

# 目录

<b>第一章</b>	<b>引言</b>	<b>4</b>
1.1	问题提出	4
1.1.1	研究背景	4
1.2	研究设计	6
1.2.1	研究内容与研究目标	6
1.2.2	研究方法与研究思路	6
1.2.3	资料获取路径	7
<b>第二章</b>	<b>指标体系构建</b>	<b>8</b>
2.1	指标体系	8
2.1.1	指标体系的构建	8
2.1.2	评价指标的数据搜集	10
2.2	模型构建	10
2.2.1	模型建立	10
2.3	模型测算	11
2.3.1	数据搜集与整理	11
2.3.2	数据标准化	11
2.3.3	模型的检验	11
2.3.4	提取主成分	14
2.3.5	计算特征向量矩阵与主成分矩阵	14
2.3.6	因子正交旋转和主因子命名解释	15
<b>第三章</b>	<b>多元线性回归分析</b>	<b>17</b>
1.1	模型介绍	17
1.1.1	回归方程的拟合优度检验	17
1.1.2	回归系数的显著性检验	17
1.1.3	适用性	18
1.2	数据验证与操作	18
1.2.1	多重共线性诊断	18
1.2.2	模型建立与结果分析	19
<b>第四章</b>	<b>聚类分析</b>	<b>21</b>

第五章 总结.....	22
5.1 主成分分析法模型优缺点分析 .....	22
5.2 分析过程总结 .....	23
参考文献.....	24
附录：原始数据表.....	24

## 第一章 引言

### 1.1 问题提出

#### 1.1.1 研究背景

随着我国经济的快速发展,人们的生活水平稳步提升,在追求物质生活的同时,也不断追求着精神层面的满足,观看电影是当下大众消遣娱乐的基本形式,好的影视作品不仅能带来欢乐,也能带来不错的票房促进影视企业的发展。据统计,2018年我国电影票房高达417亿美元,电影行业比以往任何时候都更受欢迎,每年的内地影市,都有五六百部影片上映。虽然影片的票房不能代表一切,但是票房收入则是衡量一部商业片是否成功的最重要的指标之一。了解电影票房的影响因素,有利于企业提前做好预判,有效抵御风险,因此对于电影票房的预测显得格外重要。80年代,美国的BarryLitman提出了票房收入预测模型,该模型能分析预测不同种类电影的票房价值,对之后美国电影投资界产生了颠覆性的影响。2012年我国企业也研究推出第一套票房预测系统,推动了我国影视产业的进一步发展。

在经历了前几年的行业寒冬后,中国电影业发展在近三年似乎出现了回暖迹象,迎来了新一轮的蓬勃发展。电影市场以高歌猛进的姿态,以良好的发展态势取得了一系列的斐然成就。根据国家新闻出版广电总局电影局发布的数据显示,2019年,全国电影总票房超过624亿元人民币,同比增长5.4%;国产电影票房占票房总额的64.07%;全国生产电影1037部,银幕9708块,银幕总数已达到69787块。骄人业绩的背后是居民消费水平的提升及对文化产品需求的增加,二者并驾齐驱将中国电影发展推入了快车道,与此同时,行业的逐步规范化和成熟度的提升似乎也昭示着国产电影无限的生机与潜力。相关业内人士估计,中国超过美国成为全球第一大电影市场已是指日可待的事情。在商业化浪潮席卷社会每一个角落的当下,电影市场成为资本竞相逐利的新空间。当电影的商品属性和艺术属性逐步割裂后,二者之间的空隙使得电影票房数量的多寡和电影自身质量高低常常呈现出截然不同的态势。然而,与此前国产电影高票房、低口碑的常态不同,近三年荧幕上涌现的优质大片,使票房与口碑更加趋向于吻合态势。作为“超现象级”影片,票房超过56亿元人民币的主旋律电影《战狼2》,为国产电影市

场注入了一剂强心针；以《冈仁波齐》《二十二》为代表的艺术类影片，突破重围，找到了自身的定位；《我和我的祖国》《攀登者》《中国机长》等献礼电影将宏大叙事与个体叙事巧妙融合，通过商业与文艺的交织奏响主旋律，在 2019 年 10 月 1 日的当天创下中国影史国庆档单日票房新高，实现了强口碑与高票房的双丰收。在 2017 年，我国第一部文化产业领域的法律《电影产业促进法》正式实施，我国最早于 2003 年启动电影产业立法工作，2011 年《电影产业促进法》（征求意见稿）正式公布，这部“姗姗来迟”的法律，不仅标志着电影行业规范性又迈进了一步，也彰显了电影产业发展在国民经济和社会发展规划中的重要地位。

电影孕育于文化背景下，扎根于社会生活中。资本市场力量的介入可以带来极为可观的票房数据，却无法帮助一部电影真正获得良好口碑，受众的品味和艺术欣赏力也不会轻易向高票房电影屈服，时评分网站就成为了受众喜好的最真实直接的表达渠道。自媒体时代，若选择避而不谈社交网站上受众对其的态度呈现，就无法全方位地理解或评判电影。从某种程度上来说，票房说明经济效益的成败，而评分则是艺术价值的社会认可度。票房成为业界衡量一部电影成功与否的重要标准，评分是受众对电影态度的直观反馈，二者似乎成为了决定一部电影社会声誉及影响力的全部因素。票房数据和评分数值在电影的完整公映期均会出现较为剧烈的变化，在外界看来，二者之间存在千丝万缕的纠葛，若即若离、些许“暧昧”，相互牵制又相互促进。也就是说，评分值与票房存在一定的相互影响或交织关系，且票房经济和社会文化心理共同作用于当代电影的状态。以大众娱乐方式和传播信息的有力媒介而被大众熟知的电影，作为体验型商品，其评分所展现出来的口碑评判不仅代表着受众对其的接受和认可程度，影响着购买及观看行为，更是受众对于电影本身作为艺术品抑或商品的价值阐释。在中国电影市场票房飞速攀升、互联网技术对电影产业格局和生态不断激荡的背景下，评分网站迅速扩张、其功能体验亦得到深度开发，电影网络评价体系、受众评价心理、网络水军、评分与电影质量规制、在线评分网站评价模式等议题的探讨呈现广泛而深入的态势。在电影出品方与发行方视角下，网络评分可以帮助其调整销售策略和排片流程，一定数量的票房在提供经济效益之外，又成为电影生命周期延长、艺术价值

拓展的有力保障。

由以上的分析我们可以看出影响电影票房的因素众多。作为观众，观众想了解一部电影具备哪些因素时能提供一个不错的观影体验；而作为电影投资方、导演，他们更想了解选择何种题材、在何时上映、选择哪些演员、进行何种方式的宣传能为一部电影带来效益最大化的收入。因此我们小组选取对有可能影响票房的诸多因素进行相关分析。

## 1.2. 研究设计

### 1.2.1 研究内容与研究目标

根据研究目的，我们设定研究目标为哪些因素对票房影响最明显，因此选择的因素有导演、演员、上映档期、豆瓣评分等 12 种影响因素，并对数据进行处理后进行主成分分析、多元线性回归分析和聚类分析。

### 1.2.2 研究方法与研究思路

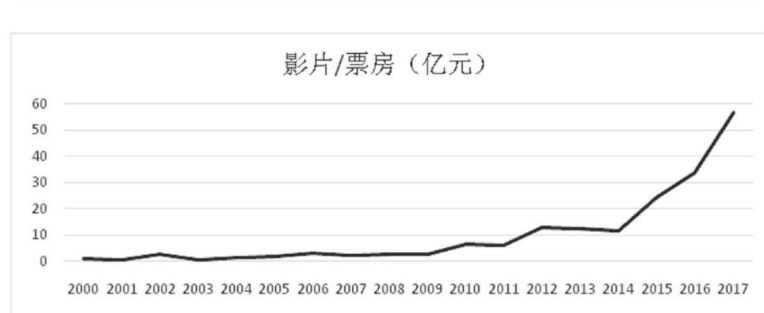
- 1) 根据研究内容，首先获得系列原始数据。
- 2) 将原始数据进行标准化处理，得到均值为 0，方差为 1 的独立正态随机变量。
- 3) 在 SPSS 上进行 KMO 和 Bartlett 检验，得出结果后剔除数值较差的公因子方差，使 KMO 值更高，为之后进行主成分分析铺垫。
- 4) 将优化后的数据进行主成分分析。
- 5) 得到初始因子的载荷矩阵（即相关系数矩阵）。
- 6) 对载荷矩阵进行因子旋转，以解释主成分含义。
- 7) 通过载荷矩阵得到主成分系数（即矩阵特征向量）。
- 8) 通过主成分系数和原始数据相乘得到主成分矩阵。
- 9) 得到的主成分矩阵进行多元线性回归分析和聚类分析。

选择主成分分析而非相关性分析的原因是：选取的各因子间存在多重贡献性，主成分分析能有效剔除这些因素；选取的部分因素对票房存在的影响不能完全通过相关性分析的数值来体现，如果仅通过相关性分析而将数值较低的因素剔除也许会对分析结果造成影响。

进行因子旋转的原因是仅通过载荷矩阵得到的主成分系数没有具体含义，通过因子旋转可以对得出的主成分进行定义、解释。

### 1.2.3 资料获取路径

图 1-1 中国 2000 年-2017 年电影票房



由图可知，中国电影票房在 2015 年开始回暖，有稳步上升的趋势。但 2020 年因受新冠疫情影响，各地电影院纷纷关停，票房因素收到巨大打击。因此我们选择 2019 年票房进行分析。因票房存在周期性的特点，我们认为选取分析一整年的票房即可有代表性。

我们选择的数据是 2019 年票房排行榜前 100 名的电影，选取的数据有：票房、上映档期、导演、演员、首日排片量、豆瓣评分等，所有数据均来源于豆瓣 APP 和猫眼专业版 APP，数据一定程度上真实可靠，有说服力。

## 第二章 指标体系构建

### 2.1 指标体系

#### 2.1.1 指标体系的构建

本文根据电影票房的理解并结合其评价指标的原则，对国内 2019 年电影总票房收入排名前一百的电影进行相关因素的分析、归纳，以影片质量、网络热度、观看意愿和时间因素四个个方面为主要内容，建立了一套评价电影票房收入水平的综合指标体系。

##### 1) 影片质量

电影质量是电影票房高低的根本和前提，而由于电影类型的不同，质量对电影票房的影响程度也不尽相同。因此影片质量部分综合考量文艺片与商业片，选取豆瓣电影评分、影片导演票房、影片主演票房作为主要内容。

##### 2) 网络热度

网络热度反映电影宣发效果和电影口碑的影响，一部电影受到越多人的讨论，就会带来更高的关注度，从而赢得更多的购买行为。本文选取豆瓣电影评分人数、百度指数峰值、微博指数峰值以及微信累计文章阅读量作为评价的主要内容。

##### 3) 观看意愿

观看意愿主要是指影响消费者想要观看某一电影的倾向或反映消费者观看某一电影的意愿，主要包括：平均票价、场均人次、电影时长、上映首日排片占比以及上映前三天累计票房。

##### 4) 时间因素

时间因素主要是指影片上映时间在黄金档时间的占比和以及上映时长，该指标能较大程度地影响票房的高低。

根据以上四个方面内容，我们选取了 14 个不同的指标，并将指标编号如下：

X1 豆瓣电影评分。豆瓣电影是中国最大的电影与评论社区，是拥有群众基础最广，相对最具公信力的大众评分网站。

X2 黄金档期时间。黄金档期时间指该电影上映时间是黄金档期的天数，黄金档期主要有：1 月 10 日-2 月 19 日为贺岁档、5 月 1 日-5 月 7 日为五一档、7 月 1 日-8 月 31 日为暑假档、10 月 1 日-10 月 7 日为国庆档，其他时间均为普通



档。

X3 上映首日排片占比。该指标反映的是该电影在上映首日全国各影院排片时段所占总排片时段的百分比。

X4 豆瓣电影评分人数。豆瓣电影是中国最大的电影与评论社区，是拥有群众基础最广，相对最具公信力的大众评分网站。

X5 影片导演近三部票房。在商业电影领域，评价一位导演所导影片剧情逻辑、剪辑水平、摄影水平等的重要标准就是其所导演的电影票房，电影票房越高，其电影质量在观众心中越为上乘。因此选取该电影之前近三部电影票房总和作为评价指标之一。

X6 影片主演近三部票房。演员票房是反映主演所出演电影的表演技术、影片质量、影片观感的重要指标。因此选取该电影之前近三部电影票房总和作为评价指标之一。

X7 电影时长。电影时长反映该电影的放映长度，该指标影响排片、放映频率、价格进而影响观看意愿。

X8 场均观影人次。场均人次反映该电影在上映期间平均每场观看电影的人次。该指标能客观反映观众的观影意愿。

X9 平均票价。平均票价反映该电影在上映期间电影票价的平均价格。价格过高会影响观影数量，价格过低会影响票房收入。

X10 微博峰值指数。新浪微博具有庞大的活跃用户，当网民希望查询其他网民对该电影的反馈时，倾向于在微博进行讨论，因此其讨论峰值指数是重要的参考指标。

X11 微信累计文章阅读量。微信是中国最大的社交平台，其文章阅读量一定程度上反映了消费者在希望了解该电影的意愿。因此选择微信累计文章阅读量能较为恰当地反映该电影的网络热度。

X12 百度峰值指数。百度作为中国最大的中文搜索引擎网站，当网民们对一部电影产生兴趣，想了解这部电影的人员信息、预告片、相关新闻等资料时，倾向于在百度上搜索该影片。因此观察影片的网络热度，选择百度峰值指数是较为合适的。

X13 上映前三天累计票房。该指标能一定程度上反映在口碑和外界因素未产生时该电影依靠前期宣发所获得的票房收入。

X14 上映时长。上映时长指从该电影上映开始至该电影最终下映的总天数。

### 2.1.2 评价指标的数据搜集

本文从豆瓣网和猫眼网查询有关数据和指标,各指标原始数据的收集请于附录查看。

## 2.2 模型构建

### 2.2.1 模型建立

本文首先采用主成分分析(Principal components analysis, PCA)对原始数据进行降维,使得在损失很少信息的前提下把多个指标转化为互不相关的主成分指标,从而消除了多重共线性使得后续的多元线性回归成为可能。

但是,要研究各个自变量对因变量变化的影响程度的现实意义,就必须知道各个主成分的含义,而主成分分析无法对各个主成分进行很好的解释。因此我们在主成分分析的基础上使用因子分析(Factor Analysis, FA)对主成分的载荷矩阵进行旋转,从而达到解决主成分分析的现实含义的障碍。

#### 1) 对原始数据进行标准化

为消除不同变量的量纲的影响,对数据进行标准化处理。假设进行主成分分析的指标变量有 $m$ 个:  $x_1, x_2, \dots, x_m$ , 共有 $n$ 个评价对象,第 $i$ 个评价对象的第 $j$ 个指标的取值为 $x_{ij}$ 。将各指标值 $x_{ij}$ 转换成标准化指标 $\tilde{x}_{ij}$ ,

$$\tilde{x}_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}, (i = 1, 2, \dots, n; j = 1, 2, \dots, m)$$

其中,  $\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij}$ ,  $s_j = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}$ , ( $j = 1, 2, \dots, m$ ), 即 $\bar{x}_j, s_j$ 分别为第 $j$ 个指标的样本均值和标准差。对应地,称 $\tilde{x}_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}$ , ( $i = 1, 2, \dots, n$ )为标准化指标向量。

经过标准化后使得各指标成为均值为0,方差为1的独立正态随机变量,其协方差矩阵为单位矩阵 $I_m$ ,即 $F \sim N(0, I_m)$

#### 2) 计算相关系数矩阵

相关系数矩阵 $R = (r_{ij})_{m \times m}$ :

$$r_{ij} = \frac{\sum_{k=1}^n \tilde{x}_{kj} \cdot \tilde{x}_{kj}}{n-1}, (i, j = 1, 2, \dots, m)$$

式中 $r_{ij} = 1, r_{ij} = r_{ji}, r_{ij}$ 是第 $i$ 个指标与第 $j$ 个指标的相关系数。

### 3) 计算特征值和特征向量及新指标变量

计算相关系数矩阵 $R$ 的特征值 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0$ , 及对应的特征向量 $u_1, u_2, \dots, u_m$ , 其中 $u_j = (u_{1j}, u_{2j}, \dots, u_{nj})^T$ , 由特征向量组成 $m$ 个新的指标变量

[illegible]

#### 4) 因子旋转

样本相关系数矩阵 $R$ 的主成分因子分析的载荷矩阵 $A$ 为

$$A = (\sqrt{\lambda_1}u_1, \sqrt{\lambda_2}u_2, \dots, \sqrt{\lambda_m}u_m)$$

由于因子及其载荷矩阵并不唯一，且正交变换是一种旋转变换，如果选取方差大的正交旋转，即将各个因子旋转到某个位置，使每个变量在旋转后的因子轴上投影向大、小两极分化，从而使每个因子中的高载荷只出现在少数的变量上。

### 2.3 模型测算

### 2.3.1 数据搜集与整理

通过查询猫眼电影、豆瓣等官方网站的统计数据来获取相应的指标数据，各指标原始数据收集请参看附录。

### 2.3.2 数据标准化

为消除指标量纲对评价结果的影响,使指标之间具有可比性

$$x_{ij}' = \frac{x_{ij} - \bar{x}_i}{\sqrt{s_i}}, (i = 1, 2, \dots, n; j = 1, 2, \dots, m)$$

其中,  $\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij}$ ,  $s_j = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}$ , ( $j = 1, 2, \dots, m$ )

经过标准化变换后，各样本的均值和方差分别为 0，1。

### 2.3.3 模型的检验

对于模型的适用性检验，一般采用 KMO 和 Bartlett 球形检验进行变量间的相关性和偏相关性检查，KMO 取值约接近于 1，变量间的相关性越强，偏相关性越弱，因子分析进行标准化变换后的检验结果如下：

表 2-1 KMO 和巴特利特检验

KMO 和巴特利特检验		
<hr/>		
KMO 取样适切性量数。		0.631
巴特利特球形度检验	近似卡方	375.115
	自由度	91
	显著性	0.000
<hr/>		

可以看到，KMO 的结果值仅接近于 0.6，因此考虑通过根据公因子方差剔除提取信息量较低的项，公因子方差如图所示：

表 2-2 公因子方差

公因子方差		
	初始	提取
X1	1.000	0.807
X2	1.000	0.737
X3	1.000	0.714
X4	1.000	0.751
X5	1.000	0.613
X6	1.000	0.303
X7	1.000	0.741
X8	1.000	0.802
X9	1.000	0.622
X10	1.000	0.639
X11	1.000	0.616
X12	1.000	0.739
X13	1.000	0.884
X14	1.000	0.456

提取方法：主成分分析法。

经过剔除公因子方差最低的三项（即 X6, X11, X14）后，再次进行 KMO 和 Bartlett 球形检验所得到的结果为：

表 2-3 KMO 和巴特利特检验

KMO 和巴特利特检验		
KMO 取样适切性量数。		0.672
巴特利特球形度检验	近似卡方	429.890
	自由度	55
	显著性	0.000

表 2-4 公因子方差

公因子方差		
	初始	提取
X1	1.000	0.855
X2	1.000	0.936
X3	1.000	0.911
X4	1.000	0.812
X5	1.000	0.813
X7	1.000	0.809
X8	1.000	0.850
X9	1.000	0.907
X10	1.000	0.874
X12	1.000	0.917
X13	1.000	0.899

提取方法：主成分分析法。

可以看到该值接近于 0.7，可以进行主成分分析。

#### 2.3.4 提取主成分

根据主成分结果可得总方差解释表，是对各变量进行主成分提取。按照因子的方差累计贡献率达到 85%以上作为主成分提取的标准，从图中可以看出前 6 各指标的累计贡献率已达到 87.128%

表 2-5 总方差解释表

成分	初始特征值			提取载荷平方和		
	总计	方差百分比	累积 %	总计	方差百分比	累积 %
1	3.887	35.332	35.332	3.887	35.332	35.332
2	1.909	17.350	52.682	1.909	17.350	52.682
3	1.217	11.066	63.748	1.217	11.066	63.748
4	1.087	9.884	73.632	1.087	9.884	73.632
5	0.921	8.375	82.007	0.921	8.375	82.007
6	0.563	5.121	87.128	0.563	5.121	87.128
7	0.444	4.039	91.167			
8	0.343	3.120	94.287			
9	0.292	2.657	96.944			
10	0.260	2.367	99.310			
11	0.076	0.690	100.000			

提取第一主成分特征值为 3.887，其贡献率为 35.332%；提取第二主成分特征值为 1.909，其贡献率为 17.35%；提取第三主成分特征值为 1.217，其贡献率为 11.066%；提取第四主成分特征值为 1.087，其贡献率为 9.884%；提取第五主成分特征值为 0.921，其贡献率为 8.357%；提取第六主成分特征值为 0.563，其贡献率为 87.128%。

#### 2.3.5 计算特征向量矩阵与主成分矩阵

根据已获得的成分矩阵，按照以下公式计算主成分系数矩阵：

$$V_i = \frac{F_i}{\sqrt{\lambda_i}} \quad (i = 1, 2, 3, 4, 5, 6)$$

得到主成分系数矩阵为：

表 2-6 主成分系数矩阵

	X1	X2	X3	X4	X5	X7	X8	X9	X10	X12	X13
U1	0.12	0.14	0.3	0.28	0.31	0.24	0.38	0.29	0.26	0.34	0.48
U2	0.57	-0.26	-0.01	0.45	-0.18	0.44	-0.22	-0.34	-0.05	0.08	-0.07
U3	0.25	0.42	0.29	-0.02	0.38	0.08	-0.14	0.03	-0.56	-0.43	0.05
U4	0.31	0.47	-0.5	0.26	-0.24	-0.35	0.36	0.01	-0.19	0.14	0
U5	-0.01	0.52	0.43	-0.07	-0.26	-0.15	-0.14	-0.53	0.37	0.07	0.05
U6	-0.13	0.21	-0.02	0.29	-0.38	0.28	0.02	0.4	0.33	-0.6	-0.09

将其作为主成分的系数，经下列公式计算即可得到主成分的得分矩阵：

$$Z_i = \sum_{j=1}^k V_{ij} \times Z_{ij} \quad (i = 1, 2, 3, 4, 5, 6)$$

### 2.3.6 因子正交旋转和主因子命名解释

表 2-7 旋转后的主成分系数矩阵

	成分					
	1	2	3	4	5	6
X1	<b>0.845</b>	-0.208	0.013	0.148	0.079	-0.261
X2	-0.011	0.114	0.141	0.001	<b>0.950</b>	-0.019
X3	0.102	0.007	<b>0.902</b>	0.041	0.154	0.249
X4	<b>0.853</b>	0.152	0.000	0.185	0.047	0.160
X5	-0.020	0.476	<b>0.661</b>	0.230	0.013	-0.309

X7	0.683	0.104	0.426	-0.073	-0.336	0.181
X8	0.110	<b>0.657</b>	0.034	<b>0.534</b>	0.303	0.171
X9	-0.045	<b>0.939</b>	0.128	0.036	0.018	0.068
X10	0.007	0.111	0.145	0.307	-0.023	<b>0.864</b>
X12	0.193	0.110	0.115	<b>0.891</b>	-0.061	0.237
X13	0.264	0.485	0.519	0.487	0.221	0.197

第一主成分（U1）在 X1（豆瓣得分）、X4（豆瓣评论数）、指标上有较大载荷，分别是 0.845、0.853，该主成分主要反映的是豆瓣网评价情况。

第二主成分（U2）在 X8（场均观影人次）、X9（平均票价）指标上有较大载荷，分别是 0.657、0.939，该主成分主要反映的是电影场均收益情况。

第三主成分（U3）在 X3（排片占比）、X5（导演前三部电影票房）指标上有较大载荷，分别是 0.902、0.661，该主成分主要反映的是影院对导演受欢迎程度的认知情况。

第四主成分（U4）在 X8（场均观影人次）、X12（百度指数峰值）指标上有较大载荷，分别是 0.534、0.891，该主成分主要反映的是观众观影与讨论倾向。

第五主成分（U5）在 X2（热门档期天数）指标上有较大载荷，分别是 0.950，该主成分主要反映的是上映档期情况。

第六主成分（U6）在 X10（微博指数峰值）指标上有较大载荷，分别是 0.864，该主成分主要反映的是微博讨论情况。



### 第三章 多元线性回归分析

#### 1.1 模型介绍

一个因变量与两个或更多的预测变量之间的联系被称为多元相关。在这种情况下做出的预测被称为多元回归。线性关系指因变量与自变量之间存在一次函数关系。多元线性回归模型如下：

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_{m-1} x_{m-1} + \varepsilon \quad (1)$$

其中  $y$  为因变量； $x_1, x_2, \dots, x_{m-1}$  为自变量； $\beta_0, \beta_1, \beta_2, \dots, \beta_{m-1}$  是  $m$  个未知参数； $\varepsilon$  是均值为零，方差为  $\sigma^2 > 0$  的不可观测的随机变量，称为误差项，并通常假定  $\varepsilon \sim N(0, \sigma^2)$ 。

多元回归模型初步建立后，是否真正解释了预测变量和因变量的关系，还要进行显著性检验。

##### 1.1.1 回归方程的拟合优度检验

判定指数  $R^2$  描述了由自变量的线性函数值所能反应的  $Y$  的总变化量比例。结果在 0-1 间，越大说明拟合效果越好，判定指数公式如下

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} = 1 - \frac{(\sum_{i=1}^m (\hat{y}^{(i)} - y^{(i)})^2)/m}{(\sum_{i=1}^m (\bar{y}^{(i)} - y^{(i)})^2)/m} = 1 - \frac{MSE(\hat{y}, y)}{Var(y)} \quad (2)$$

其中，MSE 表示均方误差，SSE 被称为残差平方和。SSR 则被称为回归和，反映了自变量的线性函数在各组观测值出取值的离差平方和；SST 则被称为总离差平方和，用于度量  $y$  自身的差异程度，即数据总的变动。

##### 1.1.2 回归系数的显著性检验

$t$  检验是分别检验回归模型中各个回归系数是否具有显著性，以便使模型中只保留哪些对因变量有显著影响的因素。检验时先计算统计量  $t_i$ ；然后根据给定的显著水平  $\alpha$ ，自由度  $n - k - 1$  查  $t$  分布表，得临界值  $t_\alpha$  或  $t_{\alpha/2}$ ，若  $t > t - \alpha$  或  $t_{\alpha/2}$ ，则回归系数  $b_i$  与 0 有显著差异，反之，则与 0 无显著差异。统计量  $t$  的计算公式为：

$$t_i = \frac{b_i}{s_y \sqrt{C_{ij}}} = \frac{b_i}{s_{bi}} \quad (3)$$

### 1.1.3 适用性

由于原始数据各指标在主成分分析中已找到相应的主成分，而主成分之间线性无关，符合自变量两两线性无关要求，因此在下列共线性诊断中，各成分的 VIF 值均接近于 1，远小于 10。

表 3-1 VIF 检验

系数<sup>a</sup>

模型		共线性统计	
		容差	VIF
1	U1	.991	1.009
	U2	.997	1.003
	U3	.995	1.005
	U4	.992	1.008
	U5	.978	1.023
	U6	.979	1.022

a. 因变量：票房

在进行主成分分析过程中，实际上也对所有原始指标进行了合并和筛选，因此在本部分回归中无需对自变量进行进一步操作。

## 1.2 数据验证与操作

### 1.2.1 多重共线性诊断

分析的基本假设，自变量之间要求相互独立的。如果某两个或多个自变量之间出现了相关性，导致参数估计出现偏差，或者标准误很大，模型失真，则成为存在多重共线性。因此在建立多元线性回归模型之前，首先要进行多重共线性诊断。

图 3-1 线性回归



SPSS 可以通过计算方差膨胀因子（VIF）来检验变量间的多重共线性，结果如下。

表 3-2 VIF 检验

模型		共线性统计	
		容差	VIF
1	U1	.991	1.009
	U2	.997	1.003
	U3	.995	1.005
	U4	.992	1.008
	U5	.978	1.023
	U6	.979	1.022

a. 因变量：票房

方差膨胀因子等于容忍度的倒数。一般情况下，VIF 的值不应该大于 5，放宽到容忍度的水平，就是不应该大于 10。由结果来看，各主成分的方差膨胀因子均接近 1，说明该数据可以通过多重共线性诊断。

1.2.2 模型建立与结果分析

由于自变量过多，选择步进的方法，此方法可以自动剔除对因变量影响较小的变量。

德宾-沃森 (Durbin-Watson) DW 用来检验回归分析中的残差项是否存在自相关 (序列相关) 现象，当 DW 值显著的接近于 0 或 4 时，则存在自相关性，而接近于 2 时，则不存在 (一阶) 自相关性。一般来说越接近 2 越好，说明自变量的自相关性越不明显，模型设计的越好。由此可见，德宾-沃森的 DW 值良好，但判定指数  $R^2$  较低，说明自变量的自相关性不明显，但整体回归方程的拟合程度一般。

表 3-3 回归拟合程度

模型摘要 <sup>c</sup>					
模型	R	R 方	调整后 R 方	标准估算的误差	德宾-沃森
1	.682 <sup>a</sup>	.465	.459	.735424198	
2	.706 <sup>b</sup>	.499	.489	.715091453	1.649

a. 预测变量: (常量), U1  
b. 预测变量: (常量), U1, U4  
c. 因变量: 票房

通过下列系数表格，我们可以得到回归方程相应的具体参数。通过 t 检验求得各成分显著性，并用步进法剔除不明显成分后，保留“豆瓣评价体系得分”和“观众对电影的偏好程度”两个成分，即  $U_1$  和  $U_4$ ，故标准化的回归模型为：

$$\begin{aligned}\text{票房} &= 0.674 * \text{豆瓣评价体系得分} + 0.186 * \text{观众对电影偏好程度} \\ &= 0.674 * U_1 + 0.186 * U_4\end{aligned}$$

第四章 聚类分析

聚类分析是根据对象的特征，按照一定的标准对研究对象进行分类。由于主成分指标的多元线性回归拟合程度不佳，故考虑在得到主成分得分矩阵的基础上进行聚类分析。同时主成分分析解决了指标间的多重共线性和数据冗余问题，增强了聚类分析的准确性。

适用 K 均值聚类，经迭代 100 次后得到最终聚类结果如下：

表 4-1 最终聚类结果

最终聚类中心					
聚类					
	1	2	3	4	5
U1	-0.589610	-0.406124	8.563699	4.900762	1.138464
U2	1.907048	-0.342330	0.623242	1.269176	-2.110862
U3	0.144756	0.093175	2.545607	-2.383314	-0.222822
U4	0.189634	-0.224057	-3.240318	1.407403	1.002162
U5	-0.410721	0.299179	-1.622840	0.673697	-1.479248
U6	0.262545	-0.113936	-0.686753	-0.153305	0.438370

表 4-2 每个聚类中的个案数目

每个聚类中的个案数目		
聚类	1	19.000
	2	67.000
	3	1.000
	4	4.000
	5	9.000
有效		100.000
缺失		0.000

经处理后得到均值的个案中心结果：

表 4-3 均值的个案中心

报告						
平均值						
个案聚类编号	U1	U2	U3	U4	U5	U6
1	-0.58960999	1.90704815	0.14475552	0.18963405	-0.41072132	0.26254510
2	-0.40612432	-0.34232985	0.09317457	-0.22405664	0.29917901	-0.11393611
3	8.56369942	0.62324163	2.54560668	-3.24031778	-1.62284022	-0.68675349
4	4.90076191	1.26917609	-2.38331424	1.40740270	0.67369668	-0.15330455
5	1.13846356	-2.11086229	-0.22282232	1.00216167	-1.47924833	0.43837048
总计	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000

结论如下：

第一类：极差的电影质量，极高的场均票房收益的电影。此类电影票房表现最差。

第二类：较差的电影质量，较低的场均票房收益，一般的观影和讨论倾向，但具有较好的档期的电影。此类电影票房表现一般。

第三类：极佳的电影质量，极佳的院线资源。此类电影票房表现最好。

第四类：极佳的电影质量，较高的场均票房收入，较好的观影讨论倾向，但院线资源较低的电影。此类电影票房表现较好

第五类：一般的电影质量，极差的场均票房收入，较好的观影和讨论倾向，较差的上映档期，较好的微博热度。此类电影票房表现波动较大，不确定性强。

## 第五章 总结

### 5.1 主成分分析法模型优缺点分析

1) 优点：

- 利用降维技术用少数几个综合变量来代替原始多个变量，这些综合变量集中了原始变量的大部分信息。
- 通过计算综合主成分函数得分，对客观经济现象进行科学评价。
- 聚类分析模型的优点就是直观，结论形式简明。

2) 缺点:

- 当主成分的因子负荷的符号有正有负时, 综合评价函数意义就不明确, 命名清晰性低。
- 聚类分析法模型优缺点分析
- 聚类分析法是指将个体(样品)或者对象(变量)按相似程度(距离远近)划分类别, 使得同一类中的元素之间的相似性比其他类的元素的相似性更强。目的在于使类间元素的同质性最大化和类与类间元素的异质性最大化。
- 其主要依据是聚到同一个数据集中的样本应该彼此相似, 而属于不同组的样本应该足够不相似。因此 K-均值法要求分析人员事先知道样品分为多少类。
- 在样本量较大时, 要获得聚类结论有一定困难。由于相似系数是根据被试的反映来建立反映被试间内在联系的指标, 而实践中有时尽管从被试反映所得出的数据中发现他们之间有紧密的关系, 但事物之间却无任何内在联系, 此时, 如果根据距离或相似系数得出聚类分析的结果, 显然是不适当的, 但是, 聚类分析模型本身却无法识别这类错误。

## 5.2 分析过程总结

通过我们的整体实验步骤和最终得出的数据结论, 与实际情况和普遍主观感受大致相符, 证明实验方向可靠。但我们注意到, 实验在主成分分析时得到 4 个主成分; 而最后进行多元线性回归分析后我们仅能保留两个因素。我们分析认为数据量不足是造成分析有误差的主要原因。从整个实验过程来看, 我们选取的模型和方法是可行的, 但由于单纯从 APP 内获得数据的方式不够高效, 数据普遍存在误差, 可选取变量不够多, 导致主成分分析(剔除过多)、结论也许与实际表现有一定偏差。在后续改进中我们希望通过各类软件、获取更真实、数据量更大的数据, 再做实验。相信能得到更为真实的数据还原。

## 参考文献

- [1]孙晓南. 基于偏回归的动漫电影票房影响因素实证研究[D]. 中北大学, 2020.
- [2]刘志新. 中国电影票房影响因素分析[J]. 合作经济与科技, 2019(17):114-116.
- [3]方薇. 网络口碑对国产艺术电影票房的影响[D]. 浙江财经大学, 2019.
- [4]叶子. 2017 年国产片网络口碑与票房的关系研究[D]. 中国电影艺术研究中心, 2018.
- [5]连聪聪, 杜仪, 张辉. 基于电影营销因素分析的电影票房预测[J]. 中国传媒大学学报(自然科学版), 2015, 22(04):26-30.

## 附录：原始数据表

片名（2019）	X1	X2	X3	X4	X5	X6	X7	X8
哪吒之魔童降世	500363	8.4	36	0.325	1607443	0	8.03829	110
流浪地球	468110	7.9	25	0.114	1666072	5.272	64.04	125
复仇者联盟 4：终局之战	424900	8.5	7	0.826	936809	43.51	34.78	181
我和我的祖国	312332	7.7	7	0.341	867537	11.03	15.86	155
中国机长	290375	6.7	7	0.299	640960	15.13	16.99	111
疯狂的外星人	221177	6.4	25	0.204	612237	0	15.86	116
飞驰人生	172624	6.9	25	0.209	796609	28.34	37.25791	98
烈火英雄	170333	6.5	30	0.36	281482	10.51	11.44322	120
少年的你	155611	8.3	0	0.429	1266477	2.37	11.21	135
速度与激情：特别行动	143454	6.3	8	0.522	221258	9.89	16.8911	137
蜘蛛侠：英雄远征	141740	7.7	63	0.537	513159	12.24	20.2216	127
扫毒 2：天地对决	131136	6	56	0.392	220966	25.41	21.28	99
大黄蜂	114945	7	50	0.502	340454	8.84	15	114
攀登者	109502	6.1	7	0.342	285916	7.58	64.04	125
惊奇队长	103496	6.9	0	0.489	450873	7.79	11.58	124
比悲伤更悲伤的故事	95775	4.8	0	0.115	154473	7.73	3.66182	106



哥斯拉 2：怪兽之王	93722	6.3	35	0.525	239584	7.88	4.89181	132
阿丽塔：战斗天使	89678	7.5	8	0.39	527742	5.51	16.2	122
银河补习班	87767	6.3	25	0.414	325459	4.21	18.98	147
狮子王	83373	7.3	23	0.317	317604	10.67	7.15	118
误杀	133311	7.6	40	0.273	711168	6.21	9.83	112
冰雪奇缘 2	82272	7.1	30	0.359	324131	9.46	7.15	103
反贪风暴 4	79858	6	0	0.362	145155	10.54	5.54341	96
叶问 4：完结篇	118170	6.9	30	0.316	259722	11.5	8.65	107
熊出没：原始时代	71708	6.6	30	0.044	19076	18.5	14.2	92
使徒行者 2：谍影行动	70192	6.3	18	0.332	173759	8.2	19.42	98
大侦探皮卡丘	64059	6.5	7	0.376	347003	5.88	11.05	104
新喜剧之王	62715	5.7	18	0.19	366858	25.3	15.58494	91
千与千寻	48821	9.4	12	0.276	1846659	5.1	2.0141	125
绿皮书	47862	8.9	0	0.171	1274690	4.1	0.857	130
白蛇：缘起	45495	7.8	30	0.124	618750	2.24	1.34345	99
最好的我们	41719	5.3	30	0.085	100893	2.45	2.02959	110
X 战警：黑凤凰	40915	5.8	10	0.369	214509	2.15	8	114
诛仙 1	40470	4.5	18	0.389	616546	5.18	0.22484	101
“大”人物	37986	6.5	28	0.199	195524	3.68	4.52501	107
何以为家	37475	9.1	7	0.14	824241	4.28		126
阿拉丁	36814	7.5	20	0.257	276330	7.49		128
驯龙高手 3	36794	7.4	0	0.302	226604	7.98	4.9186	104
来电狂响	64139	5.8	17	0.116	258628	5.48	1.0835	103
老师·好	35540	6.7	0	0.148	242690	2.89	1.887	111
终结者：黑暗命运	35529	6.9	0	0.382	150127	2.88	11.75942	128
沉睡魔咒 2	34613	6	0	0.166	155783	12.29	2.11	119
调音师	32519	8.3	0	0.094	779527	1.89		139
罗小黑战记	31501	8.2	0	0.244	445474	1.56		101

追龙 II	30906	5.5	28	0.277	128293	29.4	10.22488	103
黑衣人：全球追缉	30796	5.5	28	0.354	147451	15.4	11.91	115
雷霆沙赞！	29472	6.2	0	0.338	226334	1.12		132
决战中途岛	29398	7.7	0	0.219	202544	5.21	1.28	138
天气之子	28870	7.1	0	0.238	302820	4.45	4.45	112
勇敢者游戏 2：再战巅峰	28857	6	0	0.319	96603	15.1	21.20116	122
死侍 2：我爱我家	28680	7.2	20	0.393	376466	11.1	11.05	119
小小的愿望	27883	5.1	0	0.241	123830	12.4	4.59625	107
犯罪现场	25491	6.3	0	0.201	85716	16.2	23.23	105
双子杀手	23578	6.9	0	0.292	302482	5.45	10.85	117
密室逃生	23277	7.2	17	0.085	205945	0.7		99
名侦探柯南：紺青之拳	23198	5.9	0	0.191	134834	2.25	1.84876	110
大约在冬季	22743	5.1	0	0.234	67313	0.8	7.01212	125
两只老虎	22381	6	0	0.297	97987	0.9	7.56	96
受益人	21906	6.6	0	0.237	178720	1.15	23.9128	112
航海王：狂热行动	20494	7.6	0	0.222	89456	1.54	1.07	101
南方车站的聚会	20251	7.4	0	0.269	277403	0.89	1.4446	113
玩具总动员 4	20073	8.6	28	0.179	274532	0.78	3.51049	100
利刃出鞘	20029	8.2	0	0.118	652528	0.69	19.33806	130
一条狗的使命 2	19680	7	8	0.317	108123	0.4		108
沉默的证人	18121	5.3	12	0.231	116706	8.45	8.60909	94
一吻定情	17388	4.5	8	0.192	93276	2.46	1.79808	122
宠爱	68363	6	23	0.395	432568	5.46	26.16	108
天·火	17010	4.3	0	0.268	31564	2.56	5.28	97
海王	201307	7.6	3	0.531	1841287	26.58	13.08	143
爱宠大机密 2	15423	7	21	0.141	215876	5.4	12.78438	86
神探蒲松龄	15378	3.8	6	0.078	153265	0.98	25.92	108
只有芸知道	15970	6.4	3	0.205	218546	19.87	21.08	132

影响电影票房相关因素分析报告

小飞象	14771	6.6	0	0.17	123770	2.1	1.18	112
海上钢琴师	14375	9.3	0	0.151	2342740	0.9		165
半个喜剧	18836.8	7.4	22	0.148	656213	0.8	1385.42	111
我为你牺牲	24804.9	5.5	18	0.009	553	1.98		109
愤怒的小鸟 2	14049	7	14	0.133	74554	0.19	5.12	96
下一任：前任	13277	2.8	7	0.156	74415	0.32	1.50779	99
哆啦 A 梦：大雄的月球探险记	13141	7.3	13	0.245	63425	0.8	2.31425	111
雪人奇缘	12976	7.2	7	0.089	73221	0.8		97
星球大战：天行者崛起	14380.1	6.3	9	0.132	126757	8.69	6.65	142
小猪佩奇过大年	12499	3.9	7	0.059	27447	0.7		81
上海堡垒	12383	2.9	10	0.334	323554	4.87	26.58	107
决胜时刻	12383	6.6	7	0.162	50188	5.77	0.37325	140
廉政风云	12159	5.4	7	0.082	217548	15.4	3.57976	114
夏目友人帐	11526	8.1	0	0.198	323731	0.48		104
一个母亲的复仇	11294	6.8	0	0.125	208780	0.27		146
小 Q	11191	6.7	0	0.218	68454	0.45	4.06704	107
海市蜃楼	11088	7.7	0	0.081	417578	0.99		128
鼠胆英雄	10163	5.4	18	0.132	159464	0	3.20449	106
波西米亚狂想曲	9907	8.7	0	0.058	644747	14.1	7.34374	131
催眠·裁决	10035.6	5.8	0	0.175	91818	0.7	10.07	94
绝杀慕尼黑	9181	8.2	20	0.053	247553	0.4		120
全职高手之巅峰荣耀	8562	5.6	14	0.153	51446	0		98
霹雳娇娃	7594	6.9	0	0.223	138821	2.45		98
巨鳄风暴	7447	5.6	1	0.153	72431	0.15	16.91	87
蜘蛛侠：平行宇宙	42753	8.6	10	0.191	911541	2.45	5.64	116
被光抓走的人	7091	6.9	0	0.285	238501	0.8	38.6	131
祈祷落幕时	6764	8	0	0.098	194224	0	5.3	119
风中有朵雨做的云	6498	7.2	0	0.178	490108	0.8	15.3	124

片名（2019）	X9	X10	X11	X12	X13	X14	X15
哪吒之魔童降世	23	35		617	800290	66507.2	92
流浪地球	29	44	3814431	302	1298283	79390.3	90
复仇者联盟 4：终 局之战	23	48	3976381	24	778180	128048.3	31
我和我的祖国	35	37	18570401	1164	2089892	104354.1	100
中国机长	26	37	5073452	1300	492789	75710.4	100
疯狂的外星人	30	41	1057503	7.3	381720	93744.9	84
飞驰人生	25	41	2355089	531	270879	67917.4	94
烈火英雄	19	35	7901074	165	336991	37184	100
少年的你	16	36	16350845	442	768337	58859.1	182
速度与激情：特别 行动	15	35	1646393	661	308257	67804	98
蜘蛛侠：英雄远征	17	35	2900050	13	128155	65523.1	60
扫毒 2：天地对决	16	36	329054	6.8	191339	42297.3	98
大黄蜂	11	35	2079584	89	232360	39973.3	62
攀登者	22	37	4210728	8612.9	497	41553	100
惊奇队长	14	37	1008425	1796	533464	58294.6	95
比悲伤更悲伤的故 事	15	30	779475	5.1	896396	19295.1	79
哥斯拉 2：怪兽之 王	14	36	246071	0.59	184622	45894.5	61
阿丽塔：战斗天使	14	37	530717	516	262320	42976.6	60

银河补习班	16	34	2361310	15	247691	23208.2	99
狮子王	16	35	1811300	1659	546692	37173	100
误杀	15	33		81	209677	20373.5	250
冰雪奇缘 2	12	34	1151368	1059	137577	37445.4	100
反贪风暴 4	11	34	104653	3.1	99003	28040.1	67
叶问 4：完结篇	16	35	697593	520	216401	32406	100
熊出没：原始时代	22	39	519375	0.25	6604	23538.7	96
使徒行者 2：谍影行动	17	36	773926	3	756360	29245.1	87
大侦探皮卡丘	10	35	3346631	497	127189	27519.8	81
新喜剧之王	24	42	1585475	1189	346365	44869.5	84
千与千寻	12	31	1209474	2799	280471	19233.8	69
绿皮书	13	33	877533	29	224053	11485.4	54
白蛇：缘起	10	34	706461	4.4	95576	4081	220
最好的我们	10	32	914800	665	168423	13598.6	100
X 战警：黑凤凰	11	36	23996	303	27369	25741.6	100
诛仙 1	9	36	18760691	121	301645	27317.9	95
“大”人物	9	36	492533	154	89150	8675.8	60
何以为家	10	31		116	115716	5453.2	116
阿拉丁	10	36	1428910	436	118674	12909.7	100
驯龙高手 3	11	35	1129747	233	119168	21781	100
来电狂响	10	34	1397465	10	110018	16943.5	92
老师·好	9	34	229100	357	87928	7834.8	197
终结者：黑暗命运	11	33		188	9444	19392.7	69
沉睡魔咒 2	11	35	32625	290	798602	16018.1	84
调音师	10	32	228247	238	124762	4214.5	60
罗小黑战记	8	33	1064804	271	123036	9618.5	93
追龙 II	11	36				16132.7	100

黑衣人：全球追缉	8	36	1940302	191	60038	17590.7	40
雷霆沙赞！	8	36	278778	362	141153	19979.3	33
决战中途岛	9	32	23168	24	104657	10975.7	80
天气之子	11	30	36313	304	333396	15638	67
勇敢者游戏 2：再战巅峰	9	34	652841	0.02	81460	17364.6	98
死侍 2：我爱我家	8	32	297340	348	110186	15002.4	44
小小的愿望	9	33	3083528	414	97872	13305.1	99
犯罪现场	11	35	255110	161	541162	13212.2	99
双子杀手	10	37	16439	580	101259	14650.3	82
密室逃生	10	32	50806	86	34462	5706.6	58
名侦探柯南：绀青之拳	9	31	24313	59	60114	16022.7	95
大约在冬季	7	35	2707983	408	86211	9256.9	66
两只老虎	8	34	6251166	7.9	67927	13842.9	86
受益人	8	35	1940861	472	48214	9285	69
航海王：狂热行动	10	30		96	14924	13892.3	82
南方车站的聚会	8	35	3005960	10	180959	13642	100
玩具总动员 4	9	35	219668	272	298304	9117.2	63
利刃出鞘	10	33	316943	182	88292	9530.7	100
一条狗的使命 2	6	30	110385	0.51	56454	5036.3	98
沉默的证人	11	33	1637595	174	506970	12769.8	57
一吻定情	15	38	2093471	358	106991	11995.3	75
宠爱	31	34	4968971	2362	144686	31826.3	100
天·火	11	35	569296	50	53963	10560.2	277
海王	9	35	1902378	3051	494969	64215.2	81
爱宠大机密 2	9	33	267718	66	315578	7180.9	62
神探蒲松龄	16	42	2249906	138	61603	11085.3	34

只有芸知道	9	35	1214369	214	30786	6518.3	97
小飞象	8	35	265787	523	52065	7232	35
海上钢琴师	8	32	940618	347	187109	6344.06	54
半个喜剧	10	35	668438	250	185020	4388.6	100
我为你牺牲	32	59				2398.3	100
愤怒的小鸟 2	10	32	2026847	118	457873	6550.6	60
下一任：前任	10	34		97	41392	8899.2	35
哆啦 A 梦：大雄 的月球探险记	10	30		33	5369	9512.7	59
雪人奇缘	14	37	352126	3.7	64485	4774.7	348
星球大战：天行者 崛起	8	39		0.01	11329	6092.2	98
小猪佩奇过大年	18	42	391082	170	36494	9452.1	65
上海堡垒	10	37	9328562	1240	286018	10852.8	87
决胜时刻	9	39		113	34654	4563.8	100
廉政风云	15	43	2256997	0.32	93951	8542.2	60
夏目友人帐	9	30	1332054	141	74940	6512.1	31
一个母亲的复仇	7	30	219773	64	75939	3977.4	35
小 Q	5	33	1467786	124	146127	5713.5	90
海市蜃楼	9	32	229970	185	61409	4374.5	60
鼠胆英雄	13	35	745578	77	48330	5833.1	97
波西米亚狂想曲	12	39	570833	270	134220	4166.8	91
催眠·裁决	7	36		0.2	27355	5499.1	310
绝杀慕尼黑	9	32	22926	116	31290	1183.2	60
全职高手之巅峰荣 耀	9	35	803302	67	38597	6498.5	80
霹雳娇娃	6	32	324249	2.1	799600	5308.5	54
巨鳄风暴	7	30		20	51682	3968.6	96

影响电影票房相关因素分析报告

蜘蛛侠：平行宇宙	8	36		17	35454	17947.9	93
被光抓走的人	5	36	482632	126	103131	5869.4	100
祈祷落幕时	6	32		63	20980	2948.8	32
风中有朵雨做的云	7	35		354	94047	4493.2	57