

# 票房数据可视化报告

陈冠华

## 一、背景介绍

票房是指一部电影的商业销售情况。对票房数据的研究之所以重要，是因为票房这个指标意义非凡：它反映了电影在市场上的受欢迎程度，是判断一部商业电影成败的关键变量；另一方面，那些票房更高的电影题材会得到资方的追捧，成为未来一段时间市场的主流，因此票房也指导着制作方在题材上的选择，是投资市场的风向标。票房高低决定了一部电影最终给制作方带来的收益，因此对票房的研究有助于了解中国电影产业的盈利能力，对理解相关行业有很大帮助。

与此同时，在中国文化走向世界的大背景下，对国产影片在海外票房的研究能有效衡量电影作为文化输出的载体的效果。通过对海外电影票房占中国总票房比例的研究也能更好地了解国产电影行业本身的成熟程度和与发达国家之间的差距。

## 二、描述分析

本案例统计了数据中不同类别的电影上映的频数。从图 1 中可以看出，“爱情”类电影出现的频数最高，说明中国电影制作市场最偏爱以女性为主要受众的感情题材。“剧情”，“喜剧”，“动画”分列二，三，四位，说明在中国若全家人一同观影可供选择的“合家欢”作品数量众多。此外，“惊悚”，“动作”片出现的频数也较高，说明瞄准年轻一代的电影题材也在市场中占有相当的份额。

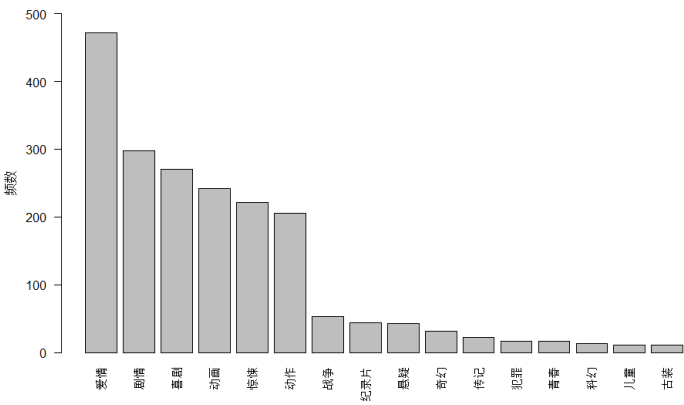


图 1 不同电影类型上映频数柱状图

从图 2 中可以看出，“科幻”，“奇幻”，“动作”电影的平均票房最高。这类电影（如阿凡达，变形金刚，碟中谍）往往在投入上耗费巨大，也因酷炫的特效和特技得到市场的狂热追捧。该结果表明“大投入，大产出”的商业模式在中国市场具备可行性。“喜剧”，“爱情”，“剧情”片的平均票房较低，可能是因为每年上映的数量过多，导致票房被同质化的竞品分流。“儿童”，“惊悚”，“传记”，“纪录”片等在平均票房上表现最差，说明这些题材在中国还未得到充分的发展和认可。“古装”题材在电影市场票房较低，可能是受到同题材电视剧冲击的影响。

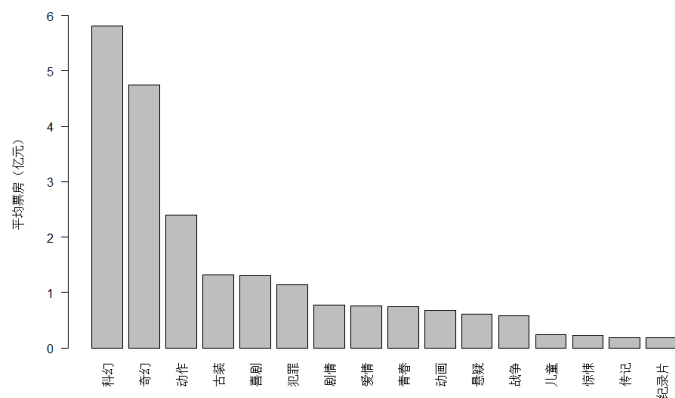


图 2 不同电影类型平均票房柱状图

### 三、投资建议

如果我是投资人，我会投资爱情题材的电影。虽然爱情片在平均票房上的表现并不亮眼，但胜在拍摄门槛低，制作周期短，票房回报稳定。反观科幻片和奇幻片，虽然各种票房榜单的冠军几乎都被他们包揽，但制作方一旦失手就要面临血本无归的窘境（比如上海堡垒投资 3.6 亿，却只收回了 1.2 亿票房）。目前国内电影工业尚不成熟，无法模仿好莱坞流水线生产商业大片，拍摄爱情题材可能尚是最稳妥的选择。

## 四、附录

```
dat = read.csv("D:/meeting/在线实习项目/数据可视化与报告撰写/TASK3 定性数据的描述统计/数据描述统计数据.csv",
              header = T)
summary(dat)#打印统计信息
dat = na.omit(dat)#删除空白值所在行
str(dat)
dat$类型 = as.factor(dat$类型)#将字符串转换成因子变量

freq = table(dat$类型)
freq.sort = sort(freq, decreasing = T)#排序
par(las = "2")#全局参数，横坐标的文字垂直排列
barplot(
  freq.sort,
  cex.names = 0.8,
  ylim = c(0, 500),
  ylab = "频数"
)

library(plyr)
dat.mean = ddply(dat, .(类型), function(x) {
  mean(x$总票房)
})#对每个类型的数据取平均值
colnames(dat.mean)[2] = "平均票房"
dat.mean = dat.mean[order(dat.mean$平均票房, decreasing = T), ]#排序
dat.mean$平均票房 = dat.mean$平均票房 / 10000#将单位化为亿元
par(las = "2")#全局参数，横坐标的文字垂直排列
barplot(
  height = dat.mean$平均票房,
  names.arg = dat.mean$类型,
  ylim = c(0, 6),
  ylab = "平均票房（亿元）"
```