

# 定量数据的描述——以小龙虾，二手房和票房数据为例

陈冠华

## 一、小龙虾销量分析

俗话说“民以食为天”，美食对于国人而言无疑是一种特别的刚需。说起美食，就不得不提这几年风头正盛的小龙虾，根据中国水产学会的报告<sup>1</sup>，自“十三五”至今围绕这一网红食品已经形成了一个庞大的产业链。该报告指出 2019 年小龙虾总产业规模已增至 4100 亿，而其中以餐饮为主的第三产业约有 2960 亿元，占 72.2%。从前景的角度看，小龙虾还能再飞一会儿。湖南，湖北等主产省都颁布了指导意见，要加大在融资和推广上的政策扶持力度。养殖技术和冷链运输环节也在发展，这意味着未来更低的原材料成本和更广阔的市场需求。

销量分析是餐饮行业中最基本的研究方法之一。在宏观层面，销量研究能把握小龙虾的市场需求，为衡量产业（尤其是餐饮行业）的发展现状提供了参考。而若将研究的范围缩小到某个区域（比如某市，某区），那么分析也能帮助相关门店了解市场需求的变化，完成竞品分析，进而据此调整经营策略，扩大营业收入。

### 1.1 销量直方图分析

从图 1 中可以发现：小龙虾销量小于 10000 份/月的门店超过 2000 家，占压倒性的比例。这个结论说明小龙虾是一个产品同质化高，需求分散，竞争充分的餐饮市场，原料与人力成本可能会成为利润的决定性因素。从右图可以发现：对销量取对数后数据分布更接近正态分布了。

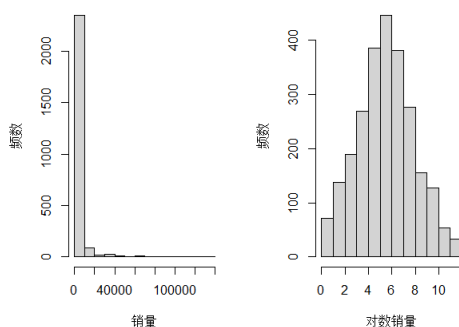


图 1 销量分布直方图

<sup>1</sup> 《中国小龙虾产业发展报告（2020）》 <http://www.boyar.cn/article/1083701.html>

## 1.2 城市平均销量柱状图分析

从图 2 中可以看出：小龙虾每月销量的平均水平依次是武汉，长沙，北京，深圳，其中武汉的月均销量超过了 7000 份。当地饮食习惯（比如是否常吃夜宵，是否嗜辣），总体消费能力以及竞品情况也许能解释这种区域性的不同。

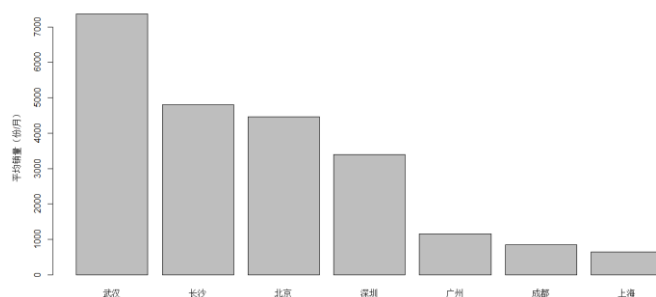


图 2 各城市平均销量柱状图

## 1.3 城市销量箱线图分析

从图 3 中可以看出：除了成都，广州，上海，其他四所城市均有销量远超平均水平的“爆款店”。将该结论与图 2 对比可以发现：爆款店是否存在与城市平均销量有一些关系。事实上，爆款店的销量抬高了平均水平，而在销量更高的地区商家更有动力进行小龙虾口味的研发，因此产生爆款的可能性更高。

对销量取对数改进极端值带来的影响，可以发现：整体上平均销量更高的地区波动程度也越大，武汉，长沙，北京，深圳的波动程度要高于广州，成都，上海。所有城市的箱长无显著差异，说明就平均水平周围 50% 的数据而言，城市与数据的离散程度关系不大。

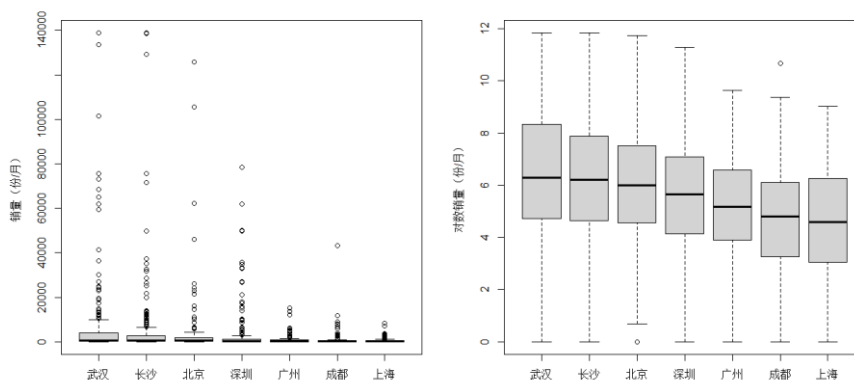


图 3 不同城市小龙虾销量对比箱线图

## 二、北京市二手房房价分析

从图 4 中可以发现：北京市学区房平均房价水平超过非学区房，波动程度也更大。这说明教育附加值对房产定价有显著影响。

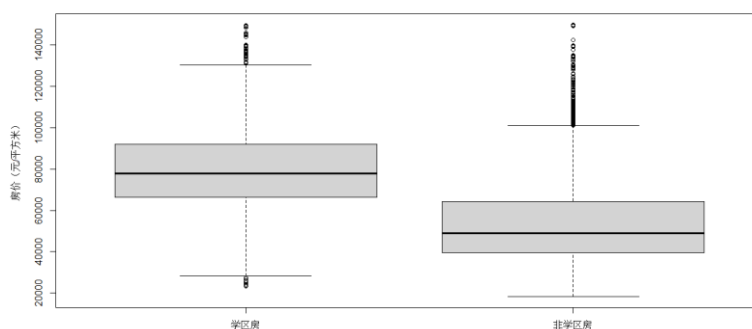


图 4 学区房房价对比箱线图

从图 5 可以看出：西城，东城，海淀，朝阳，丰台，石景山的平均房价排名顺序不受房屋是否为学区房影响。同时除了石景山以外，所有城区学区房平均房价均高于非学区房。

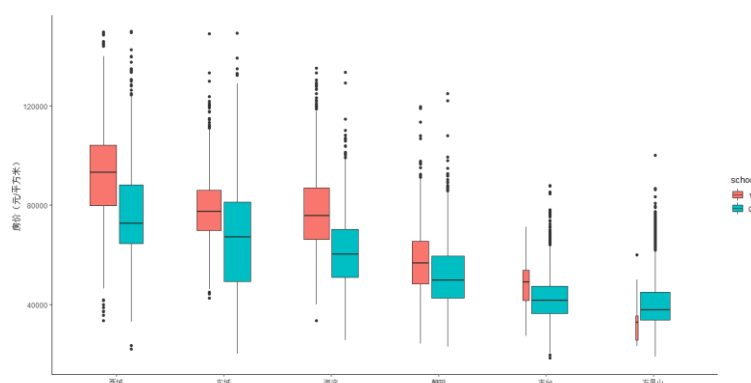


图 5 不同城区分组房价箱线图

## 三、电影票房分析

从图 6 中可以看出：奇幻，科幻，动作，古装类电影的票房平均水平最高，可能是因为这几类电影往往走的是高投入，高产出的制作思路。剧情，纪录，儿童类的影片平均票房最低，可能是因为总体制作水平还有待提高，或者市场还不够成熟。科幻，奇幻，动作，喜剧，爱情和剧情片的票房波动程度较大，背后的原因可能是高投入影片（科幻，奇幻，动作）往往风险与机遇并存，而小成本片（爱情，喜剧，剧情）一旦成为黑马也会获得高额回报。

箱线图的特点是将平均水平，波动程度和极端值都包含在一个图中，比起其他统计图（如柱状图）具有信息量的优势。

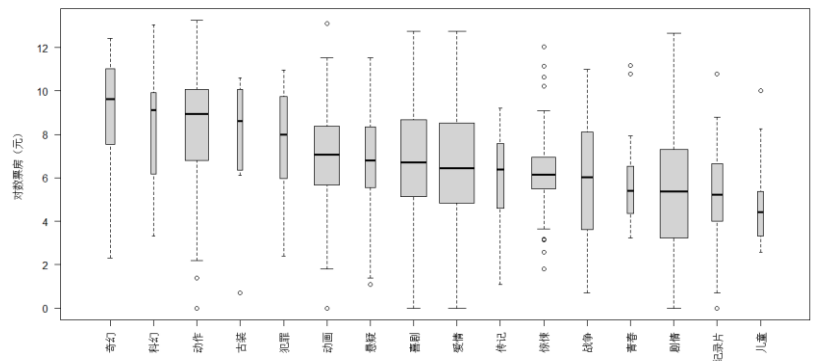


图 6 不同电影类型总票房箱线图