2 京津冀地区居民文化消费总量研究

文化消费是指用文化产品或服务来满足人们精神需求的一种消费，主要包括教育、文化娱乐、体育健身、旅游观光等方面。本研究是从三省市文化消费总量，选取北京、天津、石家庄、保定四个代表性城市居民作为调查对象，进行区域文化消费市场的定量研究，以期探究区域内居民文化消费习惯与倾向，从而有针对性引导培育文化消费市场，推动产业区域化整合及协同发展，本研究每个城市随机发放问卷200份，共发放问卷800份，经严格筛选，共收集有效问卷723份，其中北京有效问卷178份，天津有效问卷168份，石家庄有效200份，保定有效问卷177份。

2.1 京津冀各地区居民每月文化消费次数调研结果

经过对京津冀各地区居民每月文化消费次数调研，可发现各个城市均有很大比例的居民对自己每月文化消费次数没有明确次数；北京在确定每月文化消费次数中四次以上占比居多；天津、石家庄、保定在确定每月文化消费次数中二次以上占比居多。该调研结果表明，虽然京津冀地区人缘相亲，文化一脉，但是各区域文化消费情况却各有区别，且均有较大待开发潜力，所以有针对性引导培育文化消费市场对提高居民文化消费很有价值，其次北京与其他地区协同发展对京津冀共同发展有重要意义。

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 城市 | 一次占比 | 二次占比 | 三次占比 | 四次占比 | 四次以上占比 | 不一定占比 |
| 北京 | 11.24% | 11.8% | 10.67% | 2.25% | 16.29% | 47.75% |
| 天津 | 10.12% | 20.24% | 11.9% | 9.52% | 11.9% | 36.31% |
| 石家庄 | 8.0% | 17.0% | 12.5% | 10.5% | 16.0% | 36.0% |
| 保定 | 6.21% | 10.73% | 9.04% | 7.91% | 7.91% | 58.19% |

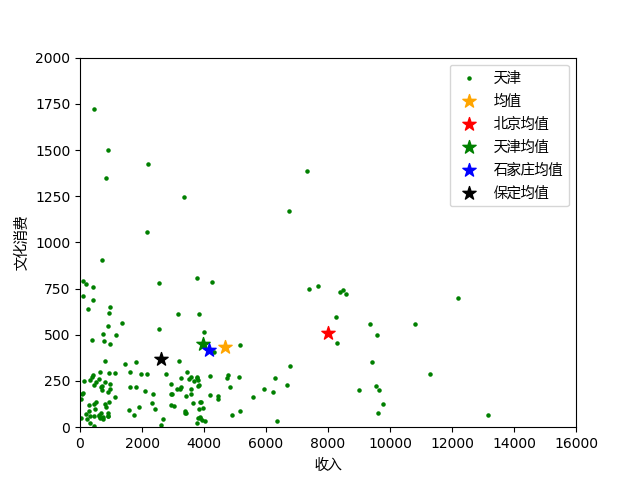
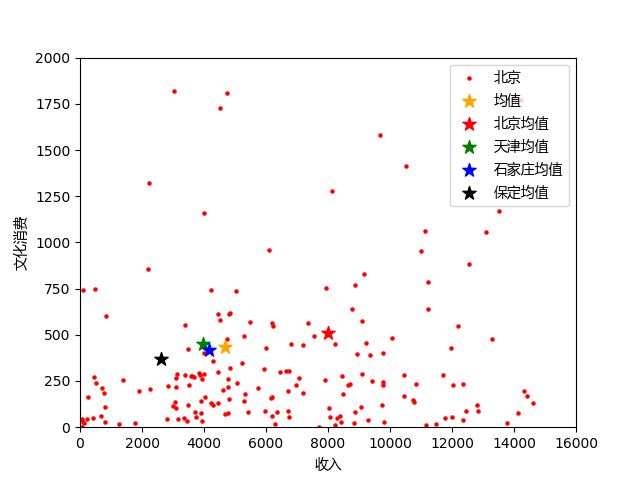
2.2 京津冀各地区居民每月文化消费资金数额调研结果

经过对京津冀各地区居民每月文化消费资金数额调研，可发现石家庄和保定居民每月文化消费资金数额在100元以下范围内占比远高于北京和天津；各个城市均有很大比例的居民的每月文化消费资金数额在100-500范围；北京和天津居民每月文化消费资金数额在500-1000范围内远高于石家庄和保定；各个城市均有很小比例的居民每月文化消费资金数额在1000以上范围；而在居民每月文化消费资金数额2000元以上的居民占比北京远高于其他三个城市。该调研结果表明，虽然京津冀地区人缘相亲，文化一脉，但是各区域文化消费情况却各有区别，且均有较大待开发潜力，所以推动产业区域化整合及协同发展对提高居民文化消费很有价值。

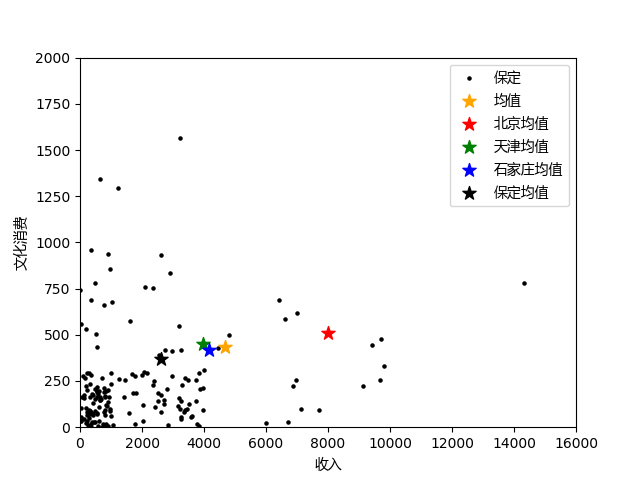
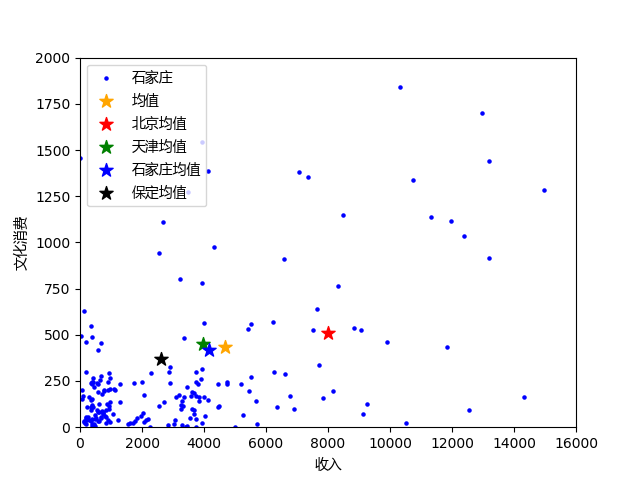
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 城市 | 100元以下占比 | 100-500元占比 | 500-1000占比 | 1000-2000占比 | 2000元以上占比 |
| 北京 | 25.84% | 39.33% | 24.72% | 6.74% | 3.37% |
| 天津 | 22.62% | 47.62% | 23.81% | 5.36% | 0.6% |
| 石家庄 | 35.5% | 42.5% | 11.5% | 9.0% | 1.5% |
| 保定 | 34.46% | 45.2% | 15.82% | 3.95% | 0.56% |

2.3 京津冀各地区居民每月文化消费总量调研结果

对于京津冀各地区居民每月文化消费总量情况，本研究采用各个区域居民文化消费支出和经济收入两个维度的散点图（下图）表示，这样可以清楚、直观地观察四座城市各个居民每月文化消费总量情况。由散点图可以看出北京市收入状况高于其他三个城市，但是文化消费状况基本和其他城市相同，说明在北京这个文化产品较为丰富的城市仍有较大的文化消费潜力；另外各区域文化消费总量分布和全部区域的均值进行比较会发现，保定表现最不乐观，几乎集中在图形左下角，超过均值的寥寥无几，如何引导该地区文化居民文化消费在该地区显得极为重要。



北京 天津



石家庄 保定

3 影响个人文化消费因素的研究

3.1 理论基础及方案设计

追求个人效用最大化的过程，也是进行属性构建、趣味区分、文化分类和社会关系再生产的过程。同时，随着社会结构变迁、价值观改变、艺术审美变化等事实因素的引入，使得定量分析法在文化消费研究中的应用就成为可能。另外，需要注意的是，中国政治及文化的不同，使得文化消费研究中“社会阶层结构”在国内研究的嫁接，更多表现于性别、年龄、职业、收入等人口统计变量上的划分，并不存在明显的政治性地位划分。

在上述理论研究的基础上，关于本研究方案的理论设计，在充分考虑到本研究的目的与数据分析的可操作性后，将主要包括的内容:各个区域影响个人文化消费因素。即：文化程度、年龄、经济收入、选择主因、获知途径。本研究还选择5个人口统计变量作为控制变量来分析它们之间存在的关联性。其中设置的具体选项如下：

1、文化程度：初中及以下、高中及中专、本科及大专、硕士研究生、博士研究生及以上；

2、年龄：18岁以下、18-30岁、30-40岁、40-50岁、50-60岁；

3、经济收入：无收入、3000元以下、3000-5000元、5000-8000元、8000-15000元、15000以上；

4、选择主因：个人兴趣、媒体宣传、优惠折扣、亲朋好友的评价、进行消费的便利程度、惯性消费；

5、获知途径：电视、网络、报刊、文化消费场所阵地广告、亲朋好友、户外广告、移动媒体（地铁、公交广告）、手机、广播、其他；

目前最有影响力的关联规则挖掘的基础算法是Apriori算法(Agrawal and Srikant, 1994)，本次分析就是采用的Apriori关联分析算法。

存在关联规则条件为:设I={i1, i2, i3, i4,……，im}是项的集合。设任务相关的数据D是数据库事务的集合，其中每个事务T是项的集合，使得T⊆I。每个事务有一个标志符，称作TID。设A是一个项集，事务T包含A当且仅当A是T的子集，形如A=>B的蕴涵式，其中A, B是I的子集，并且A∩B为空集。规则A=>B在事务集D中成立，具有支持度S，其中S是D中事务包含A∪B(即A和B两者)的占总数的百分比，它是概率P(A∪B)。规则A=>B在事务集D中具有置信度c，如果D中包含A的事务中同时也包含B的比例是c，这是条件概率P(B|A) 即:support(A=>B)=P(A∪B); confidence(A=>B)=P(B|A)o

强规则:同时满足最小支持度阈值(min-sup)和最小置信度阈值(min-conf)的规则。采用0和1之间的值表示支持度和置信度。

关联规则挖掘分为两个步骤:

(1)找出所有频繁项集(支持度不小于用户设定的阈值min-sup的项集)。令Lk表示所有频繁k一项集的集合。Apriori算法使用递推的方法找出所有的频繁项集:首先找出L1，再用L1找出L2，用L2找出L3，以此类推，直至无法找出更多的频繁项集

(2)从频繁项集中生成所有强关联规则。

Apriori算法应用广泛，可用于消费市场价格分析，猜测顾客的消费习惯，比如：

1. 湖南大学计算机与通信学院发表的《基于消费行为的Apriori算法研究》
2. 河南广播电视台的《基于Apriori算法的郑州市手机消费市场关联规则挖掘》

将关联分析算法应用在本次研究的各区域影响个人文化消费因素探究，即在用户填报的问卷当中找出上述控制变量之间的关联。

3.2 北京区域影响个人文化消费因素探究

总体状况：最小支持度：0.35，最小置信度：0.6

|  |  |
| --- | --- |
| 频繁项集 | 支持度 |
| 选择主因（个人兴趣） | 0.77 |
| 获知途径（网络） | 0.53 |
| 文化程度（本科及大专） | 0.47 |
| 年龄（18-30岁） | 0.46 |

关联规则：1、获知途径（网络）——>选择主因（个人兴趣），置信度等于0.86

2、年龄（18-30岁）——>选择主因（个人兴趣），置信度等于0.78

3、文化程度（本科及大专）——>选择主因（个人兴趣），置信度等于0.77

北京居民每月文化次数比重占比最大为四次以上，针对这类人进行因素关联分析：

最小支持度：0.35，最小置信度：0.6

|  |  |
| --- | --- |
| 频繁项集 | 支持度 |
| 选择主因（个人兴趣） | 0.93 |
| 年龄（30-40岁） | 0.52 |
| 经济收入（5000-8000元） | 0.48 |
| {经济收入（5000-8000元）=>选择主因（个人兴趣）} | 0.48 |
| {年龄（30-40岁）=>选择主因（个人兴趣）} | 0.45 |
| 文化程度(硕士研究生） | 0.41 |

关联规则：1、经济收入（5000-8000元）——>选择主因（个人兴趣），置信度等于1.00

2、文化程度(硕士研究生）——>选择主因（个人兴趣），置信度等于0.92

3、年龄（30-40岁）——>选择主因（个人兴趣），置信度等于0.87

3.3 天津区域影响个人文化消费因素探究

最小支持度：0.35，最小置信度：0.6

|  |  |
| --- | --- |
| 频繁项集 | 支持度 |
| 选择主因（个人兴趣） | 0.61 |
| 文化程度（本科及大专） | 0.55 |
| 年龄（18-30岁） | 0.50 |
| 获知途径（网络） | 0.43 |

关联规则：1、年龄（18-30岁）——>文化程度（本科及大专），置信度等于0.74

2、文化程度（本科及大专）——>选择主因（个人兴趣），置信度等于0.68

天津居民每月文化次数比重占比最大为二次以上，针对这类人进行因素关联分析：

最小支持度：0.35，最小置信度：0.6

|  |  |
| --- | --- |
| 频繁项集 | 支持度 |
| 文化程度(本科及大专） | 0.44 |
| 选择主因（个人兴趣） | 0.41 |
| 年龄（18-30岁） | 0.41 |
| 文化程度(高中及大专） | 0.35 |

关联规则：无

3.4 石家庄区域影响个人文化消费因素探究

最小支持度：0.35，最小置信度：0.6

|  |  |
| --- | --- |
| 频繁项集 | 支持度 |
| 选择主因（个人兴趣） | 0.67 |
| 文化程度（本科及大专） | 0.59 |
| 年龄（18-30岁） | 0.52 |
| 获知途径（网络） | 0.44 |
| 经济收入（3000元以下） | 0.4 |

关联规则：1、获知途径（网络）——>选择主因（个人兴趣），置信度等于0.83

1. 年龄（18-30岁）——>选择主因（个人兴趣），置信度等于0.76
2. 年龄（18-30岁）——>文化程度（本科及大专），置信度等于0.74

石家庄居民每月文化次数比重占比最大为二次以上，针对这类人进行因素关联分析：

最小支持度：0.35，最小置信度：0.6

|  |  |
| --- | --- |
| 频繁项集 | 支持度 |
| 年龄（18-30岁） | 0.62 |
| 选择主因（个人兴趣） | 0.62 |
| 文化程度（本科及大专） | 0.59 |
| 获知途径（网络） | 0.53 |
| {文化程度（本科及大专）=>年龄（18-30岁）} | 0.44 |

关联规则：1、文化程度（本科及大专）——>年龄（18-30岁），置信度等于0.75

2、获知途径（网络）——>选择主因（个人兴趣），置信度等于0.72

3、获知途径（网络）——>年龄（18-30岁），置信度等于0.72

3.3 保定区域影响个人文化消费因素探究

最小支持度：0.5，最小置信度：0.7

|  |  |
| --- | --- |
| 频繁项集 | 支持度 |
| 文化程度（本科及大专） | 0.73 |
| 选择主因（个人兴趣） | 0.72 |
| 年龄（18-30岁） | 0.63 |

关联规则： 1、年龄（18-30岁）——>文化程度（本科及大专），置信度等于0.93

2、文化程度（本科及大专）——>选择主因（个人兴趣），置信度等于0.72

保定居民每月文化次数比重占比最大为二次以上，针对这类人进行因素关联分析：

最小支持度：0.5，最小置信度：0.7

|  |  |
| --- | --- |
| 频繁项集 | 支持度 |
| 文化程度（本科及大专） | 0.84 |
| 年龄（18-30岁） | 0.63 |
| {文化程度（本科及大专）=>年龄（18-30岁）} | 0.63 |
| 选择主因（个人兴趣） | 0.58 |
| 经济收入（无收入） | 0.53 |

关联规则：1、年龄（18-30岁）——>文化程度（本科及大专），置信度等于0.1

2、选择主因（个人兴趣）——>文化程度（本科及大专），置信度等于0.91