

“钟灵毓秀 曲水流觞”

传统文化类综艺节目传播特征及传播效果分析

世人若被明日累，春去秋来老将至。

朝看水东流，暮看日西坠。

百年明日能几何？请君听我明日歌。



《明日歌》
〔明〕钱福

摘要

为了响应落实中办、国办联合下发的《关于实施中华优秀传统文化传承发展工程的意见》以及十九大报告“推动中华优秀传统文化创造性转化、创新性发展”的要求，过去的一年里我国的传统文化类综艺节目呈井喷之势。在当下娱乐综艺占据荧屏的时刻，《国家宝藏》、《中国诗词大会》等传统文化类综艺节目如一股“清流”，让大众重温中国传统文化，受到了观众们的一致好评。传统文化类综艺节目作为新型综艺模式在寓教于乐、传承传统文化方面显现出了极为可观的发展潜力。本文就传统文化类综艺节目这一话题提出两个问题：一是现阶段传统文化类综艺节目的发展现状如何？二是传统文化类综艺节目的传播特征和对传统文化的传播效果如何？针对这两个问题，本次调查采用了线下问卷调查及评论文本挖掘来获取数据，使用统计分析方法与大数据挖掘算法展开了研究。

调查实施过程中，我们进行了预调查（P6）和正式调查（P7）两轮调查。在正式调查中，我们以武汉市居民总体为调查对象，采用分层抽样与非严格的 π PS四阶段抽样相结合，将武汉市的所有区域按行政分区编号，再按街道划分，对于抽中的街道，前往相应居委会抽取居民。预调查问卷各项检验均通过，问卷结构合理。正式调查过程中共发放问卷 1200 份，回收有效问卷 994 份，有效回收率 82.83%，达到了最低样本量 921 份的要求。正式调查数据均通过了信度、效度、游程检验和独立性检验，表明此次调查具有较高的随机性并且大部分调查项目具有显著的相关关系，调查结果真实可靠。

我们运用描述性统计分析（P8）得到现阶段传统文化类综艺节目的发展现状；运用 Ridit 检验（P12）分析观众对节目的关注点和节目的闪光点；使用二元选择模型（P12）探究了传统文化类综艺节目观众的特征；运用轮廓分析探究节目对消除传统文化相关行业歧视的作用（P15）；运用结构方程模型（P16）探究了节目对文化传承的影响效果，有利于电视台“对症下药”，做进一步的改进分析，从而加强对传统文化的传播。同时使用文本挖掘技术（P19）分析了观众在网络社交平台上对节目的评价，对现状分析做进一步的研究和补充。通过抓取用户在微博平台上发表的关于四大热门节目数据，使用中文分词与 LDA 主题模型等文本挖掘方法，实现对文本数据的情感倾向性判断（P20）以及所隐藏信息的挖掘与分析，分析出观众对文化类综艺节目的好评度、关注点以及建议改进点。通过对数据的分析得出以下结论（P22）：

从文化类综艺节目传播特征方面来说，“公众人物+传统文化”模式对节目的传播起到了非常强的作用；观众特征影响显著，高学历家庭观众观看意愿较高。

从文化类综艺节目传播效果方面来说，节目降低了行业歧视，相关从业者更受尊重；传统文化领域受众趋于年轻化；节目特征显著，观众文化传承热爱及民族自豪感提高；节目成为了一种新的文化教育形式；文物维护及推广经费增加，从业人员待遇提升；观众二次传播意愿较高，掀起全民献策热潮。

关键词：传统文化类综艺节目，非严格的 π ps 抽样，二元选择模型，文本挖掘，Ridit 检验，轮廓分析，结构方程模型

目 录

一、引言	1
(一) 研究背景	1
(二) 研究思路	2
(三) 研究特色与创新	2
二、调查方案	3
(一) 调查目的与对象	3
(二) 抽样调查基本步骤	3
三、调查过程与数据处理	5
(一) 调查实施	5
(二) 预调查数据检验	6
(三) 正式调查数据检验	8
四、文化类综艺节目传播特征分析	8
(一) 基于描述性统计的节目传播对象分析	8
(二) 基于 Ridit 检验法的节目传播内容分析	13
(三) 二元选择模型下观众异质性分类集群分析	13
五、文化类综艺节目影响效果分析	14
(一) 基于描述性统计的居民行为及年龄层变化影响效果	14
(二) 基于轮廓分析的行业认知变化影响效果	16
(三) 基于结构方程模型的文化传承效果分析	17
六、基于文本挖掘的观众反响分析	20
(一) 评论数据处理与分词	20
(二) 基于情感词典的情感倾向性分析	21
(三) 基于 LDA 模型的主题分类	22
(四) 建议型评论分析	23
七、结论与建议	23
(一) 研究结论	23
(二) 建议	25
参考文献	26
附录一	29
附录二	33
附录三	41
附录四	46
附录五	54
附录六	56

一、引言

(一) 研究背景

文化类综艺节目响应落实了十九大报告《推动中华优秀传统文化创造性转化、创新性发展》的号召，倡导利用多种媒介平台，坚持创造性转化和创新性发展，深入挖掘中华文化的价值内涵。



图 1-1 国内各综艺收视率及豆瓣评分¹

如图 1-1 所示，收视率上，文化类综艺节目《中国诗词大会》已经赶超娱乐综艺类节目《歌手 2018》。从豆瓣评分上可以看出，娱乐综艺类节目远远不及文化类综艺节目。据此，文化类综艺节目已经具有能与娱乐综艺类节目抗衡的水平。

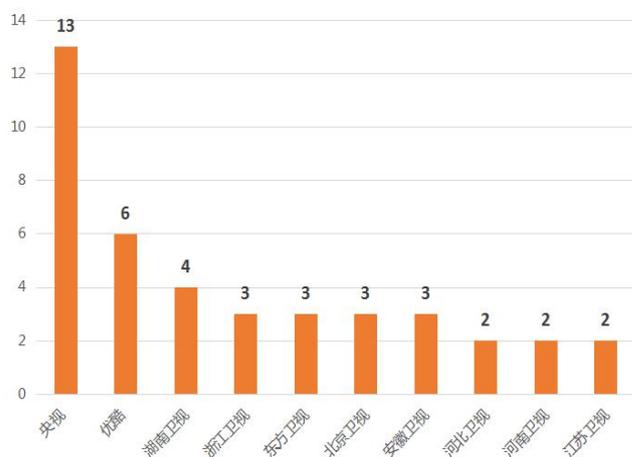


图 1-2 2017 年文化类节目播出数量²

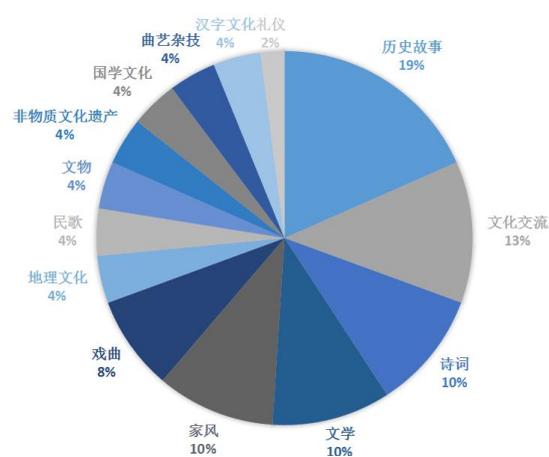


图 1-3 题材份额²

如图 1-2 和 1-3 所示，在过去的一年里传统文化类综艺节目呈现井喷之势，仅央视一家电视台就有 13 个相关类别节目，题材份额涉猎广泛，其中历史故事以 19% 的比例占据头位。

¹ 数据来源：豆瓣评分

² 数据来源：央视索福瑞

(二) 研究思路

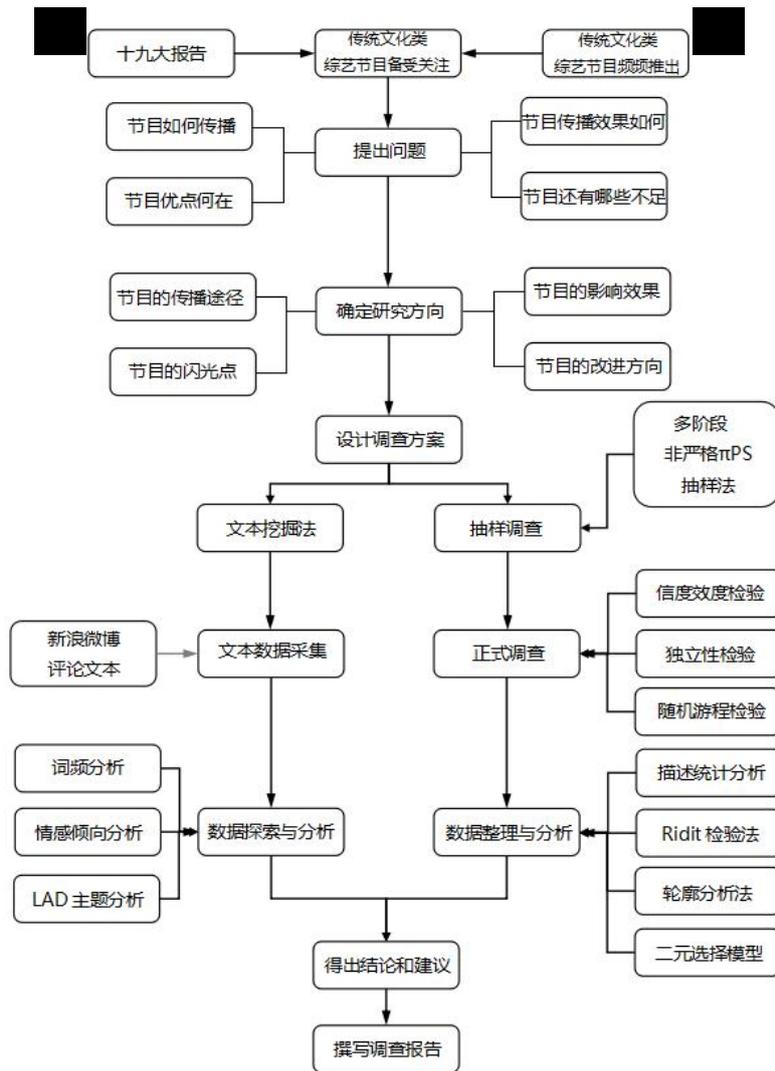


图 1-4 研究思路图

(三) 研究特色与创新

1. 基于大数据的文本挖掘

通过抓取新浪微博社交平台所有可获取的观众评价数据，通过中文分词、构建 LDA 主题模型等文本挖掘方法，从评论数据中分析出观众对文化类综艺节目的好评度、关注点以及抱怨点。

2. 实地电子问卷调研提高数据质量

我们的调研过程中采用“云调研电子问卷”——以智能客户端为展现形式，调研过程中只需要将平板电脑给被调查者填写。

3.调查内容新颖

作为近几年才出现的节目，文化类综艺节目还在不断发展更迭中，课题具有一定的研究价值。

4.研究方法丰富

除了描述性统计分析对观众群体和非观众群体进行定性分析，我们还运用Ridit 检验法，分析观众对文化类综艺节目的不同重视因素及其闪光点，并借助Logit 二元选择模型分析影响居民观看文化类综艺节目的显著性影响因素，用结构方程模型分析节目对文化传承的影响效果，另外还运用了轮廓分析和文本挖掘方法对相关方面进行分析。

二、调查方案

（一）调查目的与对象

1.调查目的

第一，搜集武汉市居民对于传统文化类综艺节目的观看情况，并搜集调查对象的基本信息，分析文化类综艺节目的观看情况和观众特征。

第二，搜集了解观众对于节目因素的重视程度以及对于节目闪光点的判断，分析节目传播特点。

第三，搜集居民观看节目前后的行为变化以及对于相关从业人员的印象变化，分析节目传播效果。

第四，搜集观众节目的建议，结合观众特征，提出节目的改进建议。

2.调查对象与单位

调查对象：武汉市居民总体

调查单位：每一个武汉市居民

3.调查项目

本次调查着眼于对传统文化类节目的观看情况、传播特点以及影响效果的研究。主要包含个人信息、节目观看情况、节目传播效果、观众偏好、观众特质五个方面。对于这五个层次，我们设计了相对应的调查项目以进行考量。(具体问卷编码表见附录三)

（二）抽样调查基本步骤

1.样本量的确定

本文采取多阶段 $n=5$ 的非严格的 πPS 抽样法进行抽样（抽样方式和方法详见附录一“抽样方式和方法”）。样本量的确定是抽样调查的基础，故首先应确定必要样本量。

如果估计的总体中具有某种属性的单位占比为 p ，可以得到近似样本量为

$$n_0 = t^2 Q / \gamma^2 P$$

其中， n_0 为近似样本量， γ 为相对允许误差， t 为一定置信度下的临界值。

由于我们采用的是复杂抽样，故所需样本量应为

$$n' = n_0 \times Deff$$

我们取 $p=0.5$ ，取相对允许误差不超过 5%，取 $Deff=1.5$ 可计算出最佳样本量为 $n'=921$ ，假定样本回收率为 80%，拟发放问卷 1200 份。

2. 编制抽样框

在非严格的 πPS 不等概率抽样法中，耶茨-格伦迪方法^[7]是简单、高效的方法，其样本抽取过程如表 2-1，抽取江岸区、武昌区、汉阳区、洪山区、江夏区。

表 2-1 耶茨-格伦迪方法抽取过程表

i	行政区	抽取第一个样本单位的代码	抽取第二个样本单位的代码	抽取第三个样本单位的代码	抽取第四个样本单位的代码	抽取第五个样本单位的代码
1	江岸区*	711084	711084	711084	711084	711084*
2	武昌区*	1076733*	-	-	-	-
3	江汉区	486676	486676	486676	486676	486676
4	硚口区	527593	527593	527593	527593	527593
5	汉阳区*	576536	576536	576536	576536*	-
6	青山区	438525	438525	438525	438525	438525
7	洪山区*	936139	936139	936139*	-	-
8	东西湖区	283103	283103	283103	283103	283103
9	汉南区	112889	112889	112889	112889	112889
10	蔡甸区	454810	454810	454810	454810	454810
11	江夏区*	588943	588943*	-	-	-
12	黄陂区	1121602	1121602	1121602	1121602	1121602

13	新洲区	958484	958484	958484	958484	958484
$\sum M_i$		8273117	7196384	6607441	5671302	5094766

由于

$$\sum M_i = M_0 = 8273117$$

所以在[1,8273117]中生成第一个随机数，得到 721003，即武昌区入样；由于 $M_0 - M_2 = 7196384$

所以在[1,7196384]中生成第二个随机数，得到 5016221，即江夏区入样；同理再进行一次操作，得到随机数 3476551，即洪山区入样。

以此类推，我们在武汉 13 个行政区中，抽取出了 5 个行政区（武昌区、江夏区、洪山区、汉阳区、江岸区）进行入样。（抽样框详见附录二表 2-2）

3.设计调查问卷

第一部分包括被访者的性别、年龄、最高学历、职业、收视偏好等信息；第二部分通过分类，对观看过文化类综艺节目的被访者和没有观看过文化类综艺节目的被访者分别进行调查，了解被访者对于节目的认知、偏好、了解渠道等情况。

4.预抽样调查

在正式调查前，先抽取少量样本进行预调查，从而避免在正式调查中发生错误和遗漏导致更大损失。我们抽取 200 个样本进行预调查，在预调查中可以对回收问卷的数据进行信度与效度检验，检验均显著。（详见 P6 “预调查数据检验”）

5.正式抽样调查

根据表 2-2 中抽样框在第四阶段抽样框内进行系统抽样。首先对所有抽中的二十七社区居民户从 1 至 3178351 进行编号。由抽样数 1000 可计算出间距为 318 后，以此进行正式调查。（详见 P7 “正式调查数据检验”）

三、调查过程与数据处理

（一）调查实施

1.调查时间安排

整个调查工作分 5 个环节，分别为前期准备、预调查、正式调查、数据的处理与分析、按照分析结果撰写调查报告。具体时间安排如图 3-1 所示：



图 3-1 调查时间安排图

2. 调查实施进度

我们将调查全过程进一步细化为 8 个阶段，并明确规定了每一阶段的起止时间、项目内容和人员分工。（具体调查实施进度表计划内容详见附录二表 3-1）

3. 抽样效果检验与控制

在本次调查中，我们采取的是多阶段 $n=5$ 的非严格的 πPS 抽样法。为判断抽样效果的好坏，控制抽样估计的精度，我们对样本设计的效果进行了测定，计算本次抽样调查的设计效果系数 $Deff$ 。

记 y_1, y_2, \dots, y_n 为按抽样中顺序排列的样本单位的指标值，相应的 Z 值为 z_1, z_2, \dots, z_n ，令

$$\begin{cases} t_1 = y_1 / z_1 \\ t_2 = y_1 + y_2 / z_2(1 - z_1) \\ t_n = y_1 + y_2 + \dots + y_{n-1} + y_n / z_n(1 - z_1 - z_2 - \dots - z_{n-1}) \end{cases}$$

则拉齐估计量^[7]为

$$\hat{Y}_{Raj} = \frac{1}{n} \sum_{i=1}^n t_i$$

它是总体总值 Y 的无偏估计量，其方差 $V(\hat{Y}_{Raj})$ 的无偏估计量为

$$V(\hat{Y}_{Raj}) = \frac{1}{n(n-1)} \sum_{i=1}^n (t_i - \hat{Y}_{Raj})^2$$

容易计算无放回简单随机抽样的估计量方差为：

$$v_{srs}(\bar{y}) = \frac{1-f}{n-1} s^2$$

因此设计效果系数应为

$$Deff = \frac{V(\bar{Y}_{Raj})}{v_{srs}(\bar{y})}$$

通过样本数据的收集与整理，我们可以计算出总方差，得到 $Deff=1.498$ ，因此我们在四阶段的简单随机抽样的基础上设计效果好很多。

（二）预调查数据检验

在数据检验开始之前，先对数据进行处理，首先进行数据收集与核查。将所

有问卷进行编号并存档，对深入访谈的结果进行记录和整理，核查问卷的完整性和准确性。接着，进行数据编码及录入（具体项目编码表详见附录四）把问卷对应的编码录入计算机中。最后，进行数据预处理。对于存在超出范围、有极端值、逻辑不一致以及空缺率在 15%以上问卷，进行剔除。

在预调查过程中，在洪山区和江夏区得到有效问卷 181 份，有效回收率为 90.5%。对问卷进行检验，依据结果进行相应的修改，提高有效性和准确性。

1. 区分度检验

以观众对文化类综艺节目中各因素满意度评分 Q17 为鉴别项目对用户进行鉴别，高分组下限得分为 4 分，低分组上线得分为 3 分。（具体各题项高低组占比详见附录二表 3-2）

确定分组后，分别计算两组在单个题目上的均分，用样本均值差异性 t 检验对每个问题进行检验。结果发现，p 值均小于 0.05，说明题目设计合理，被调查单位具有鉴别性，预调查项目通过区分度检验。因此，问卷通过了检验，可以用于正式调查。

2. 信度检验

信度是指测验结果的一致性、稳定性及可靠性，系数越大，说明测量的可信度越大。

通过对问卷 Q17 的项目进行 Cronbach α 系数检验，信度系数为 0.831，说明问卷结构和题项设计科学合理。（检验结果详见附录二表 3-3）

折半信度^[8]是指在测验后将测验项目分成相等的两组（两半），然后计算两项项目分之间的相关，系数越高表示信度高。根据问卷中 Q17 的项目得到折半系数为 0.763，系数较高，说明信度较好。（检验结果详见附录二表 3-4）

通过 Cronbach α 信度系数和折半系数的双重检验，可以发现，该问卷的信度检验结果较好，问卷结构合理。

3. 效度检验

效度是指调查的度量标准反映某一概念的真正含义的程度。

内容效度指的是测量所选择项目与测量目标和要求之间是否适合，采用单项与总和和相关分析法获得评价结果。根据 Q17 问题进行检验，发现 p 值均小于 0.05，各因子与总量表得分之间的相关性很显著，表明量表具有较好的内部一致性。（检验结果详见附录二表 3-5）

结构效度是指测量结果体现出来的某种结构与测值之间的对应程度。采用 KMO (Kaiser-Meyer-Olkin) 检验统计量来检验，KMO 系数为 0.728，P 值为 0.000，因此问卷的结构设计较好。意味着预调查所采用的问卷能够达到本次调查的目的，问卷具有结构效度。（检验结果详见附录二表 3-6）

（三）正式调查数据检验

1.信度检验

针对量表问题 Q17, Cronbach α 信度系数和折半信度系数分别为 0.845 和 0.767, 说明问卷分类合理, 问卷信度结果高, 量表内在一致性高。

2.效度检验

内容效度方面, 各项因子与总量表得分都很显著, 说明问卷 Q14 具有良好的内容效度。KMO 系数为 0.805, P 值为 0.000, 因此问卷的结构设计较好, 意味着问卷具有结构效度。

3.随机性检验

单样本变量值的随机性检验是对某变量的取值是否随机进行的检验, 也称为游程检验 (Run 过程), 根据样本标志排列所形成的游程的多少进行判断的检验。

当显著性水平为 0.05 时, 有 $-1.96 < Z = 0.55 < 1.96$, 不拒绝原假设 H_0 , 认为样本数据的变量值出现是随机的。接着, 我们还对其余的分类变量进行随机性检验, 结果表明大部分数据的变量值出现都是随机的, 不违反随机性, 可以认为本次调查问卷的结构和说服力较高。

4.独立性检验

独立性检验也称为卡方检验 (ChiSquare), 是根据次数资料判断两类因子彼此相关或相互独立的假设检验。

我们以问题“文化类综艺节目有哪些闪光点”与年龄阶段是否相关为例, 首先做出两变量的列联表。(列联表详见附录二表 3-7)

利用 SPSS 计算得到 $\chi^2=352.681$, 对应的 P 值小于显著性水平 $\alpha=0.05$, 拒绝原假设, 认为观众眼中的文化类综艺节目的闪光点与年龄阶段之间存在一定的相关关系, 不同年龄阶层的人的眼中节目的闪光点存在出入。另外, 我们还验证了其他变量之间是相互独立, 得到的结果显示大部分变量两两之间都存在一定的相关关系。

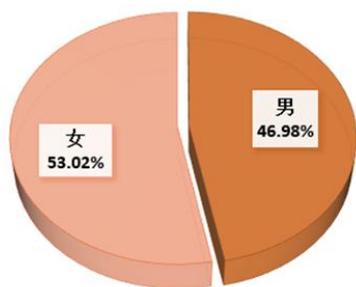
综上所述, 对正式数据进行相应的处理之后, 通过信度检验、效度检验、随机性检验和独立性检验的结果均表明本次调查是成功的。

四、文化类综艺节目传播特征分析

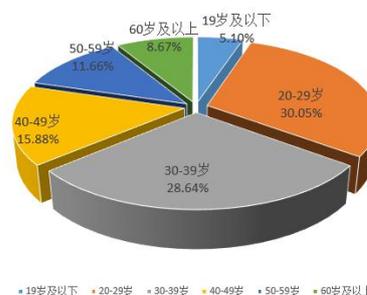
（一）基于描述性统计的节目传播对象分析

1.观众基本信息描述

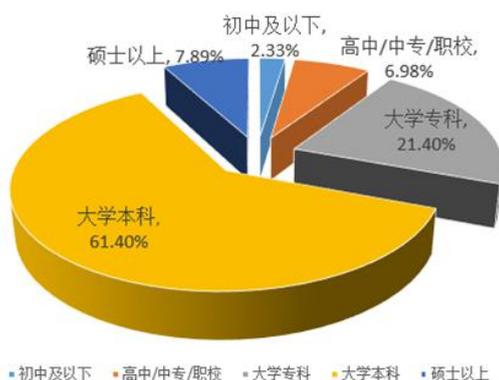
文化类综艺节目的观众中，性别分布如图 a，男性占比为 46.98%，女性占比为 53.02%，性别差异不大；年龄分布如图 b，50 岁以下观众占比 79.67%，其中 20-29 岁多达 30.05%，19 岁及以下仅占 5.10%；学历分布如图 c，拥有专科以上学历占比 90.69%，其中本科学历占 64.40%，硕士学历占 7.89%，专科学历占 21.40%；职业分布如图 d，个体经营者或承包商、农林牧渔劳动者、退休人士占比很少，其他职业分布适中。



a.文化类综艺节目观众性别分布



b.文化类综艺节目观众年龄分布



c.文化类综艺节目观众学历分布



d.文化类综艺节目观众职业分布

图 4-1 观众基本概况

如图 4-2，在传统文化认知来源这块，有 67.64% 的被调查观众选择了考纲要求，家庭教育与文化节目占比持平，分别为 42.21% 和 40.70%，这说明除了考纲之外，文化节目在被调查者的传统文化认知来源构成中占有相对重要的地位。

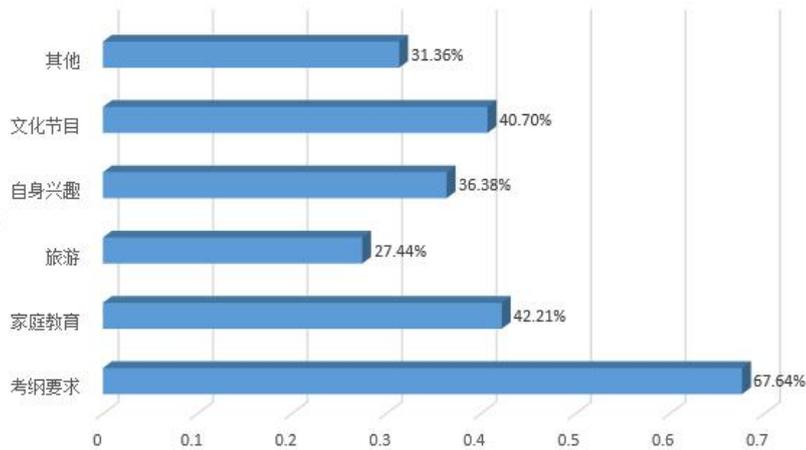


图 4-2 观众传统文化认知来源分布

如图 4-3 所示，横坐标表示观众籍贯地区，纵坐标表示观众喜爱的地区电视台。气泡大小表示某个地区观众选择某个电视台的频数。可以看出，红色斜对角虚线处的气泡相对较大，说明某个地区的人对该地区的电视台有偏好。红色水平虚线表示喜爱中央电视台的观众籍贯来源，可以发现籍贯为北方地区的气泡相对较大，说明籍贯为北方的观众相对于南方的观众更喜爱看中央电视台。

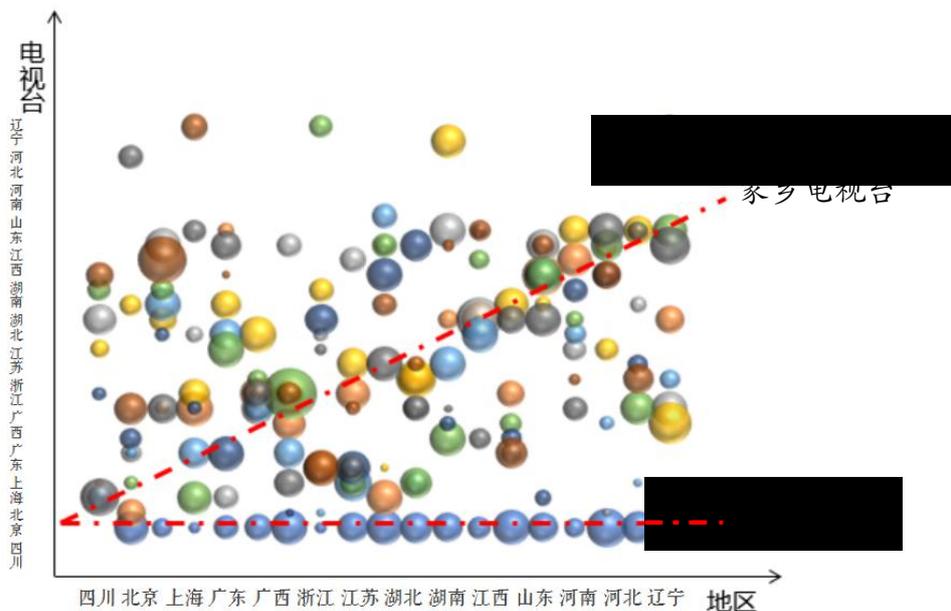


图 4-3 观众籍贯与所喜爱电视台关系图

如图 4-4，被调查者中，通过微博/朋友圈/公众号推送而得知文化类综艺节目的人最多，占比 57.69%，朋友/亲人/同学告知这一渠道位居第二，占比 41.03%。

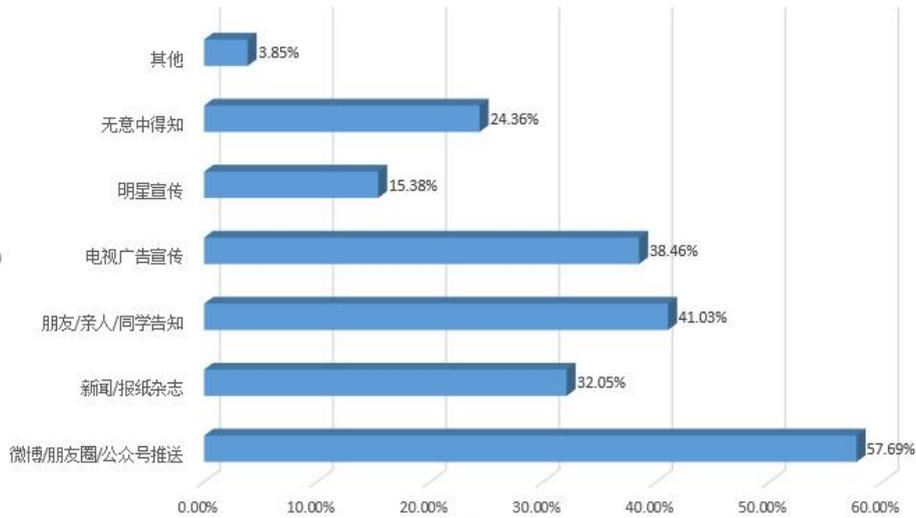


图 4-4 文化类综艺节目认知渠道分类

2. 观众观看情况描述

如图 4-5，被调查者中有 82.05% 的人最近一年有观看过文化类综艺节目，节目已经具有良好的观众认知基础；观众观看频率如图 4-6，每周观看时间少于 1 小时的观众最多，达到了 44.87%。

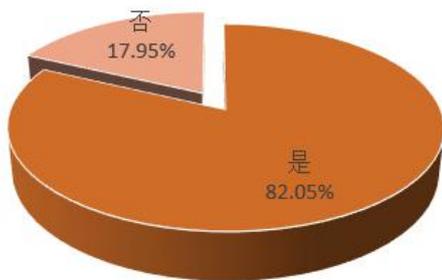


图 4-5 文化类综艺节目观看率

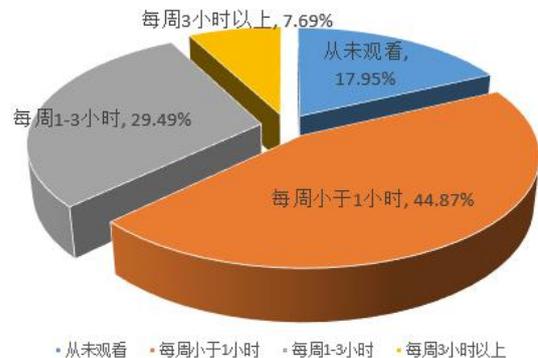


图 4-6 观众观看频率

如图 4-7，现有的文化类综艺节目中，最受欢迎的文化类综艺节目是《中国诗词大会》，占比 61.54%，其次是《国家宝藏》，占比 29.49%，《朗读者》占比 17.95%。；如图 4-8，44.46% 的被调查者认为文化类综艺节目的作用是仅供娱乐。让我们出乎意料的是，有 40.01% 的被调查者认为文化类综艺节目的作用是寓教于乐，15.53% 认为文化类综艺节目的作用还可以是传承文化底蕴。

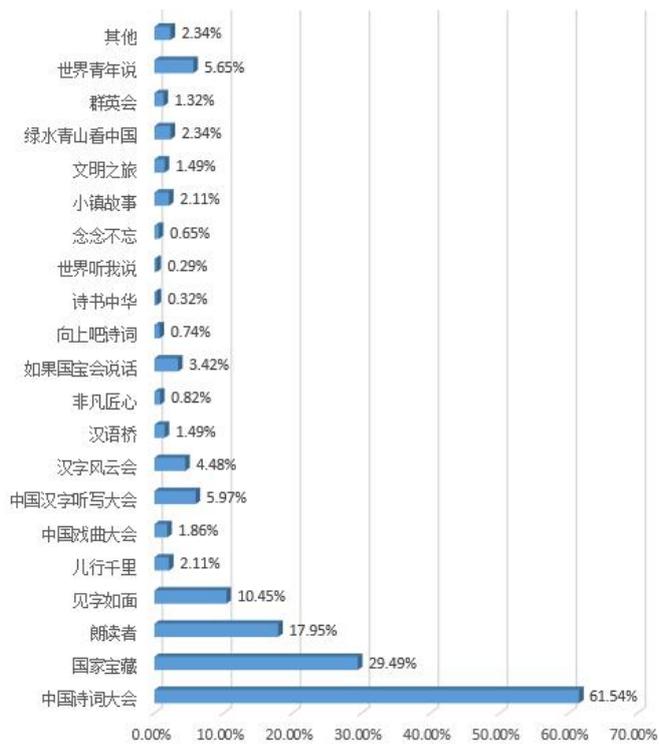


图 4-7 观众喜爱程度分布图

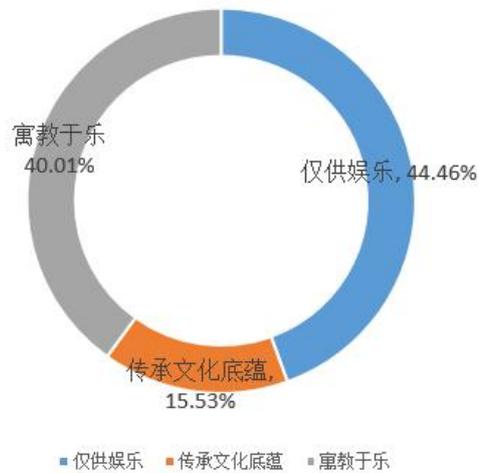


图 4-8 观众对节目看法情况

(二) 基于 Ridit 检验法的节目传播内容分析

运用 Ridit 检验法，分别计算观众对文化类综艺节目的不同重视因素、节目的不同闪光点 Ridit 得分，从而比较不同节目不同因素在观众心中的重视程度、比较节目不同闪光点的好评程度。在此，我们只展示正项得分项。

1.对观众考虑因素的分析

利用 R 软件，我们可以得到两次 Ridit 检验的结果如表 4-1 和表 4-2 所示：

表 4-1 Ridit 得分表 1

		RiditScore	Confidenceinterval		W	P-value
观众观看文化类综艺节目时可能考虑的因素	电视台影响力	0.3994	0.3272	0.4716	36.01	2.877e-07
	节目形式	0.5317	0.4595	0.6039		
	内容质量	0.6731	0.6010	0.7453		
	舞台效果	0.4639	0.3917	0.5361		
	播出时间段	0.4319	0.3597	0.5040		

由此可见，对于一档文化类综艺节目，观众最为重视的影响因素时内容质量，其次是节目形式；舞台效果对观众观看意愿影响较小；除此之外，观众并不重视电视台影响力和播出时间段；

2.对节目闪光点的分析

表 4-2 Ridit 得分表 2

		RiditScore	Confidenceinterval		W	P-value
节目的闪光点	较传统类节目更具趣味性	0.4551	0.4027	0.5076	23.92	8.282e-05
	使观众更有参与感	0.5753	0.5228	0.6277		
	明星入驻	0.5458	0.4933	0.5983		
	内容专业性	0.4182	0.3657	0.4707		
	对文化传承有积极作用	0.5056	0.4531	0.5581		

大部分观众对于现有的文化类综艺节目中逐渐增强的互动性较为满意，除此之外，明星的入驻和文化传承积极作用也是观众喜爱该类综艺节目的原因之一；

(三) 二元选择模型下观众异质性分类集群分析

从观众特征出发，通过分类集群，归纳出观看节目的群体特质。考虑到是否观看节目只涉及两种回答，是个二分类变量，建立二元选择模型。设事件发生的概率为 P_i ，观看过节目设置为 1，未观看过节目设置为 0，构建 Logit 模型。

$$\ln \frac{P_i}{1-P_i} = \sum P_i x_i (i=1,2,3,\dots)$$

$$\begin{aligned} \ln\left(\frac{P_i}{1-P_i}\right) = & \alpha_0 + \alpha_1 Male + \alpha_2 Age20_29 + \alpha_3 Age30_39 + \alpha_4 Age40_49 \\ & + \alpha_6 Age60 + \alpha_7 High + \alpha_8 Colleague + \alpha_9 Undergraduate \\ & + \alpha_{10} Postgraduate + \alpha_{11} Job1 + \alpha_{12} Job2 + \alpha_{13} Job3 + \alpha_{14} Job4 + \alpha_{15} Job5 \\ & + \alpha_{16} Job6 + \alpha_{17} Job7 + \alpha_{18} Job8 + \alpha_{19} Job9 + \alpha_{20} Kid1 + \alpha_{21} Kid2 + \alpha_{22} Kid3 \\ & + \alpha_{23} Source1 + \alpha_{24} Source2 + \alpha_{25} Source3 + \alpha_{26} Source4 + \alpha_{27} South \\ & + \alpha_{28} Middle + \alpha_{29} North + \alpha_{30} Northwest + \alpha_{31} Southwest + \alpha_{32} Northeast \end{aligned}$$

根据观众的特征分类，Logit 模型变量设定及公式中各个变量如表 4-3 所示（见附录二表 4-3 “二元选择模型变量”）。

我们可以得到，40-49 岁的居民观看意愿减弱，而硕士及以上学历、家中只有一个孩子的居民观看意愿显著增强，其余变量对观看意愿无显著影响，观看比例为 80.91%，回归显著的因子结果见表 4-4（详细结果见附录二表 4-4）。

$$P_i = \frac{e^{\alpha_0 + \alpha_1 x_1 + \dots + \alpha_{31} x_{31}}}{1 + e^{\alpha_0 + \alpha_1 x_1 + \dots + \alpha_{31} x_{31}}} = \frac{e^{\alpha_0 + \alpha_4 Age40_49 + \alpha_{10} Postgraduate + \alpha_{19} Kid1}}{1 + e^{\alpha_0 + \alpha_4 Age40_49 + \alpha_{10} Postgraduate + \alpha_{19} Kid1}} = 0.8091$$

表 4-4 Logit 模型回归显著结果

变量	系数	标准差	Wald 统计量	P 值
Age40_49	-0.6303	0.2416	-2.61	0.012**
Postgraduate	0.7076	0.2293	3.09	0.003***
Kid1	0.4530	0.2525	1.79	0.079*
Cons	0.9138	0.2460	3.71	0.001***
R2=0.4260				

五、文化类综艺节目影响效果分析

（一）基于描述性统计的居民行为及年龄层变化影响效果

1. 居民行为

在问卷中，我们将节目对于观众行为的影响效果按照行为难度，分为四个等级，级数越高，表示影响效果更强。四个等级分别为：1. 是否在日常生活中讨论与文化类综艺节目相关的话题；2. 是否会查阅相关资料；3. 是否会游历文化古迹；

4. 是否会参与传统文化相关行业或相关活动。

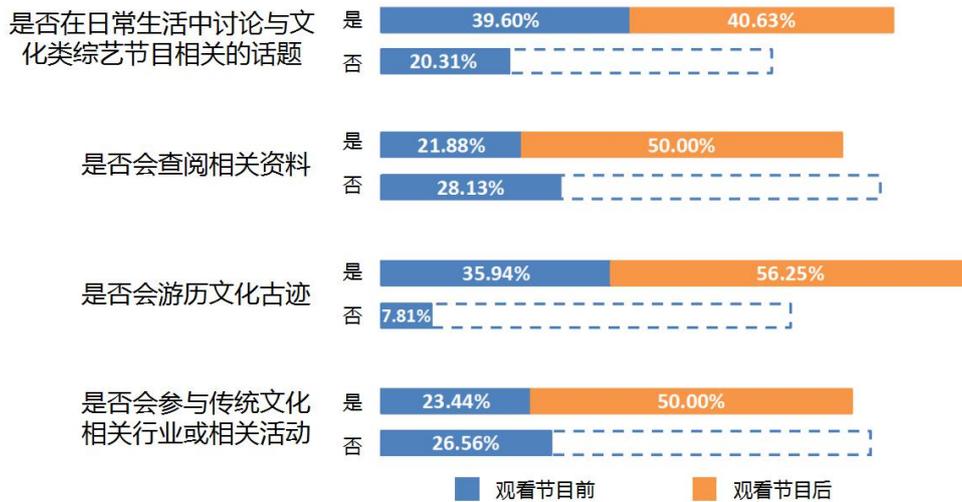


图 5-1 居民行为的影响效果分析图

如图 5-1 所示，在节目播出前，有 39.60%的居民会在生活中讨论与文化类综艺节目相关的话题；仅有 21.88%的居民会自主查阅传统文化相关资料，观众在节目播出前极不愿意查阅相关资料；35.94%愿意游历古迹；仅有 23.44%愿意参与相关活动/行业。而节目播出后，讨论节目的比例达到了 80.23%，增加了 40.63%；查阅相关资料的人数比例提升到了 71.88%，增加了 50.00%；愿意游历古迹的比例上升到了 92.19%，提升了 56.25%，于此同时只有 7.81%的居民仍不愿意游历文化古迹；愿意参与相关活动/行业的比例上升到了 73.44%，提升了 50.00%。即认为节目的播出极大地提升了人们对相关活动/行业的关注度，更多人愿意参与其中。

2. 观众二次传播意愿

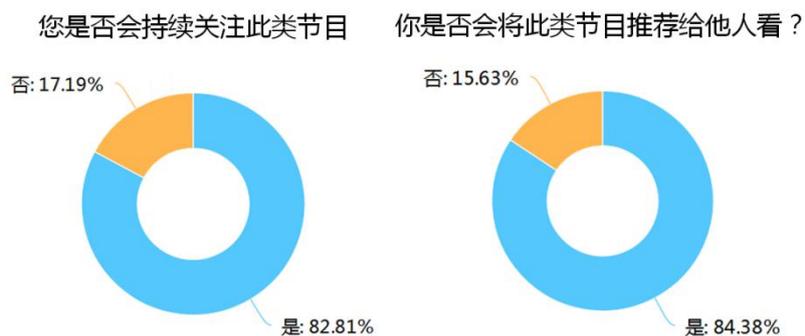


图 5-2 观众二次传播意愿的影响分析图

如图 5-2，82.81%的观众在观看节目后愿意继续关注此类节目，用户粘性极高，利于节目的继续发展；84.38%的人愿意讲节目推荐给他人，二次传播效果极好，利于节目在群众中的自行普及传播。

3.观众年龄层结构分析

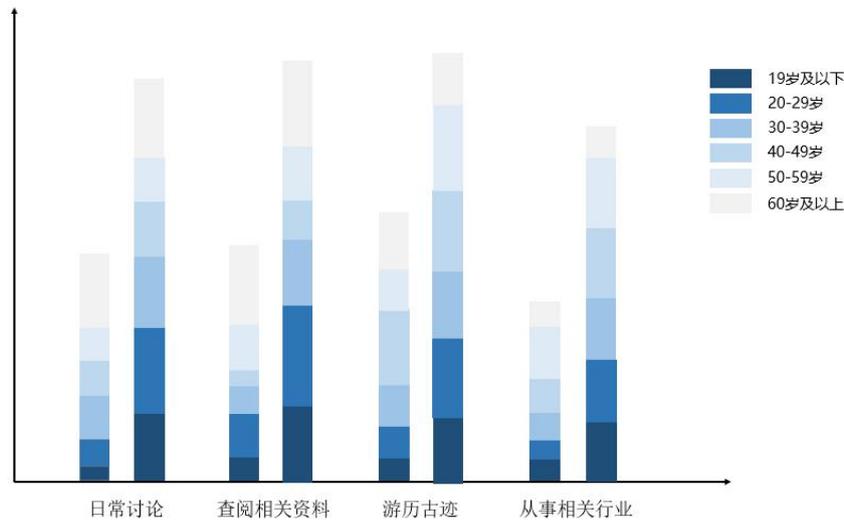


图 5-3 观看节目前后各项活动年龄层分布

如图 5-3，每项活动的左边柱状图表示观看此类节目前会进行该项活动的人年龄层分布，右边柱状图表示观看此类节目后进行该项活动的人年龄层分布。

观看节目前，日常讨论的年龄层主要分布在 50-59 岁和 60 岁及以上这两个区间，分别占比 19.33%和 15.97%；查阅相关资料的年龄层主要分布在 30-39 岁和 40-49 岁这两个区间，分别占比 19.33%和 15.97%；游历古迹的年龄层主要分布在 40-49 岁和 50-59 岁两个区间，分别占比 23.45%和 28.97%；从事相关行业的年龄层主要分布在 40-49 岁和 50-59 岁两个区间，占比分别为 18.18%与 28.79%。在观看节目后，日常讨论的年龄层主要分布在 20-29 岁及 60 岁及以上这两个区间，占比分别为 21.36%、17.96%；查阅相关资料的年龄层主要分布在 20-29 岁及 30-39 岁这两个区间，占比分别为 21.36%、17.96%；游历古迹的年龄层主要分布在 20-29 岁与 50-59 岁两个区间，分别占比 18.49%、19.75%；从事相关行业的 20-29 岁年龄层占比从 11.36%上升为 17.60%。

（二）基于轮廓分析的行业认知变化影响效果

图 5-4 为利用 spss 软件进行轮廓分析的结果图，横坐标为各行业标签，从左到右分别为非主流、有前途、低收入、温文尔雅、腐朽无能、铁饭碗、福利好待遇高、知识渊博、无私、辛苦劳累；纵坐标从 1 到 7 度量了居民对于此标签的认同的程度，认同程度越高得分越高；显示了居民在观看节目前后，对传统文化从业者的看法（蓝线为观看前，绿线为观看后）。节目播出来，居民对传统文化从业人员的待遇印象有显著改观，标签得分变化见表 5-2。（具体检验结果详见

附录二表 5-1)

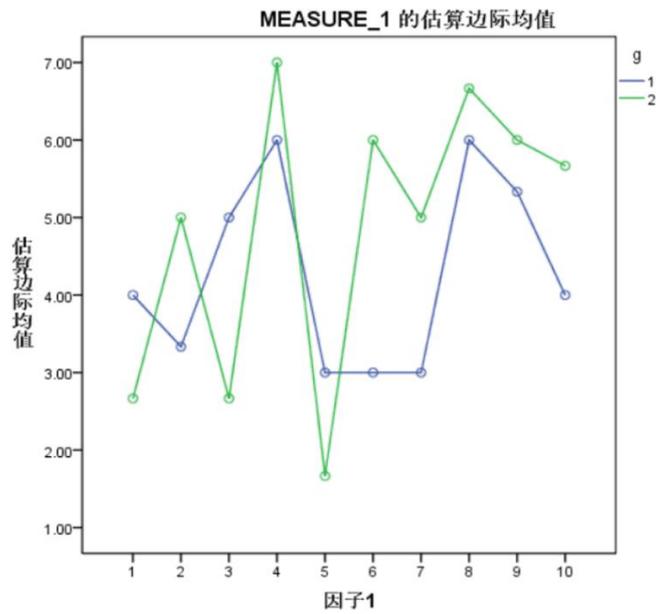


图 5-4 居民对传统文化行业看法的轮廓分析图

表 5-2 行业标签得分情况表

行业标签	观看节目前平均分	观看节目后平均分
非主流	4.02	2.71
有前途	3.35	5.01
低收入	4.99	2.73
温文尔雅	5.98	6.97
腐朽无能	3.06	1.68
铁饭碗	3.06	5.94
福利好待遇高	3.06	4.89
知识渊博	5.64	6.43
无私	5.27	5.75
辛苦劳累	4.12	5.41

(三) 基于结构方程模型的文化传承效果分析

1. 结构方程模型的构建

我们运用结构方程模型探究了民族自豪感、文化热爱程度、电视台影响力、节目趣味程度、舞台效果、播出时间段、观众参与程度、明星入驻和节目内容专业程度共9个变量之间的关系。文化热爱程度以及民族自豪感2个因子为潜变量，其对应的7个观测变量为显变量。方程采用极大似然估计法进行估计，采用多种

方法来考量模型的拟合优度。

本文提出假设如下：

- H1：电视台影响力对民族自豪感有正向影响
- H2：趣味程度对民族自豪感有正向影响
- H3：舞台效果对民族自豪感有正向影响
- H4：播出时间段对民族自豪感有正向影响
- H5：观众参与程度对民族自豪感有正向影响
- H6：明星入驻对民族自豪感有正向影响
- H7：内容专业程度对民族自豪感有正向影响
- H8：电视台影响力对文化传承热爱程度有正向影响
- H9：趣味程度对文化传承热爱程度有正向影响
- H10：舞台效果对文化传承热爱程度有正向影响
- H11 播出时间段对文化传承热爱程度有正向影响
- H12：观众参与程度对文化传承热爱程度有正向影响
- H13：明星入驻对文化传承热爱程度有正向影响
- H14：内容专业程度对文化传承热爱程度有正向影响

本文研究概念框架如图 5-5 所示。

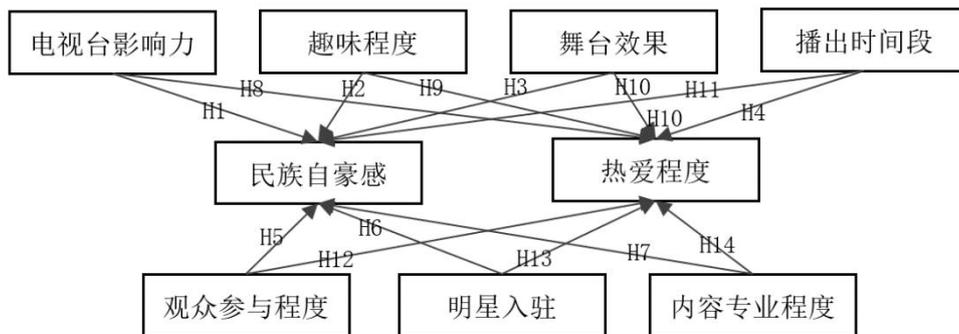


图 5-5 文化类综艺节目传承效果概念框架

2. 传承效果因素测量

基于建立的结构方程模型，通过 AMOS 软件得到如下图 5-6 的结果。

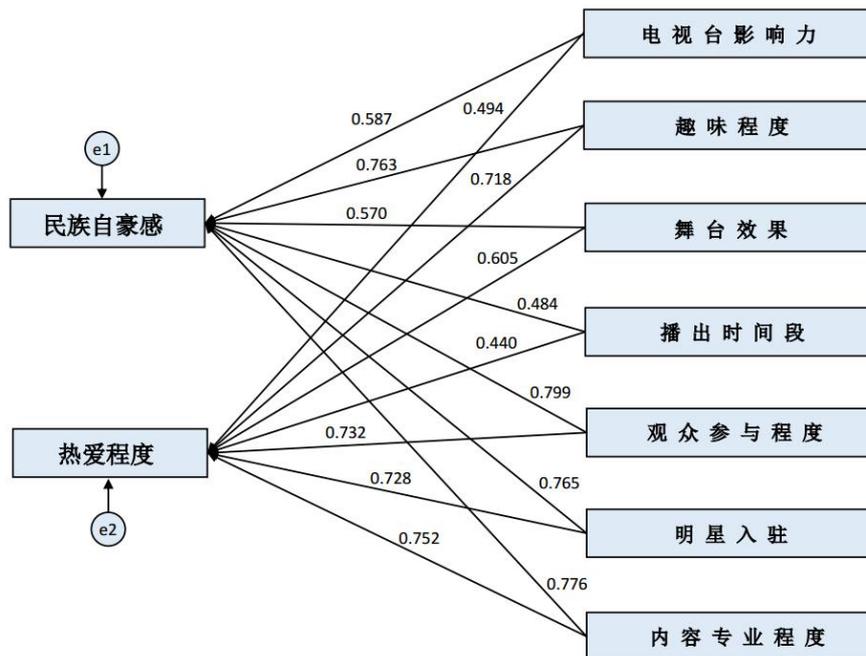


图 5-6 结构方程模型图

模型检验与拟合优度结果见表 5-3

表 5-3 修正模型检验与拟合优度结果

Model	CMIN	P	CMIN/DF	CFI	NFI	IFI	RMSEA
Default model	38.369	0.000	1.668	0.922	0.938	0.943	0.011

表 5-3 中显示，模型中卡方统计量对应的 P 值为 0.000，表现为不显著。卡方与自由度之比小于 2。此外，CFI、NFI、IFI 的值均接近于 1，RMSEA 值小于 0.02。以上指标均满足模型检验与拟合优度的要求，说明模型的拟合效果非常好。

如表 5-4 所示，可见在 5% 的置信水平下，不接受原假设 H3、H4、H10、H11，说明舞台效果和播出时间段对民族自豪感和热爱程度均有显著的正向影响。

对于民族自豪感而言，观众参与度的影响最大，系数为 0.799，其次为内容专业程度，系数为 0.776，另外电视台影响力、趣味程度、明星入驻均对民族自豪感产生显著的正向影响。

对于文化传承热爱程度而言，内容专业程度的影响最大，系数为 0.752，其次为观众参与程度，系数为 0.732，另外电视台影响力、趣味程度、明星入驻均对文化传承热爱产生显著的正向影响。

表 5-4 SEM 假设检验结果

研究假设	标准化系数	结论
H1: 电视台影响力对民族自豪感有正向影响	0.587	接受
H2: 趣味程度对民族自豪感有正向影响	0.763	接受
H3: 舞台效果对民族自豪感有正向影响	0.570	不接受

续表 5-4 SEM 假设检验结果

H4: 播出时间段对民族自豪感有正向影响	0.484	不接受
H5: 观众参与程度对民族自豪感有正向影响	0.799	接受
H6: 明星入驻对民族自豪感有正向影响	0.765	接受
H7: 内容专业程度对民族自豪感有正向影响	0.776	接受
H8: 电视台影响力对文化传承热爱程度有正向影响	0.494	接受
H9: 趣味程度对文化传承热爱程度有正向影响	0.718	接受
H10: 舞台效果对文化传承热爱程度有正向影响	0.605	不接受
H11: 播出时间段对文化传承热爱程度有正向影响	0.440	不接受
H12: 观众参与程度对文化传承热爱程度有正向影响	0.732	接受
H13: 明星入驻对文化传承热爱程度有正向影响	0.728	接受
H14: 内容专业程度对文化传承热爱程度有正向影响	0.752	接受

六、基于文本挖掘的观众反响分析

本节拟通过抓取用户在微博平台上发表的关于四大热门节目(通过预调查问卷得出,分别为《国家宝藏》、《中国诗词大会》、《朗读者》和《经典咏流传》)的数据,使用情感倾向分析与 LDA 主题模型等文本挖掘方法,以期准确的从网络社交平台文本数据中分析出观众对文化类综艺节目的好评度、关注点以及建议改进点,从而对此类节目的核心竞争力、优缺点进行评价。

(一) 评论数据处理与分词

通过 R 进行数据采集,成功抓取评论 258408 条。在采集数据的过程中,海量的原始数据中存在需剔除的数据,因此我们先对原始数据进行了清洗。本文的分词工作是使用 R 软件的 Rwordseg 包来完成的,(具体分词过程见附录一“数据清洗与分词”)将样本数据进行中文分词后,提取出高频词画出词云统计图 6-1:



图 6-1 词云统计图

如上所示，词频越高，词的字号就越大。《国家宝藏》中，观众的关注点主要集在“文化”、“守护”上面，感受有“喜欢”、“期待”等；《中国诗词大会》中，出现的高频词有“中华”、“光芒”等积极词汇，感受大多是“厉害”、“自豪”等正面情绪；《朗读者》中，高频词有“传承”、“能量”等；《经典咏流传》中，“传唱”、“中华力量”等出现频率较高。总体而言，高频词汇都是好评偏多，而“用心”、“融合”、“创新”等词可以看出观众对这些综艺节目模式的认可。

另外，我们发现明星相关的词汇几乎在每个节目中的出现率都是较高的，在国家宝藏中有“陈晓”、“演员”；朗读者中有“胡歌”、“董卿”；经典咏流传的明星效应最为突出，有频率最高的“张杰”、“毛晓彤”、“任嘉伦”，由此可以看出明星嘉宾和知名主持人对一档综艺节目有非常大的影响。

（二）基于情感词典的情感倾向性分析

本文使用的是谭松波老师的正向、逆向情感词典，作为判断文本情感倾向、情感极性的参照，进行情感倾向性分析的方法如下述公式：

$$P(di) = \begin{cases} 1(\text{sumPos}(di) > \text{SumNeg}(di)) \\ -1(\text{sumPos}(di) < \text{SumNeg}(di)) \\ 0(\text{sumPos}(di) = \text{SumNeg}(di)) \end{cases}$$

上式中 $P(di)$ 表示文本 di 的情感极性值， $\text{sumPos}(di)$ 表示文本 di 中包含的正面情感词个数， $\text{sumNeg}(di)$ 表示文本 di 中包含的负面情感词个数。对经过预处理的文本评论进行判别分析后，结论如下：

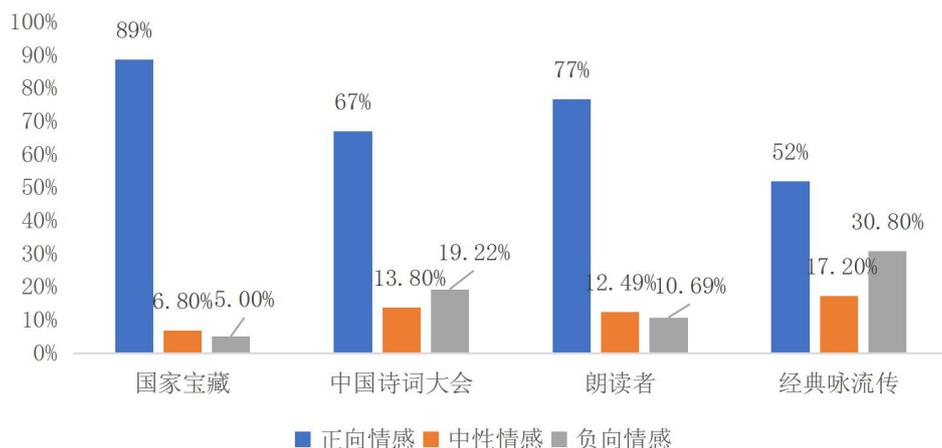


图 6-2 文本情感分析结果

如图 6-2 所示，可以看出四个样本都是正向情感偏多，其中国家宝藏和朗读者的正向情绪所占比例尤其突出，分别达到了 89% 和 77%，而中国诗词大会和经典咏流传的负向情感分别达到了 19.22% 和 30.80%。尽管负向情感所占比例总体而言没有正向情感多，但我们对于负向情感的出现更应该关注。

（三）基于 LDA 模型的主题分类

潜在狄利克雷分配 (Latent Dirichlet Allocation, LDA) 是用在一系列文档中发现抽象主题的一种统计模型，认为每一篇文档的每一个词都是通过“一定的概率选择了某个主题，并从这个主题中以一定的概率选择了某个词语”，经过 LDA 主题分析后，评论文本被聚成 3 个主题，每个主题下生成 5 个最有可能出现的词语以及相应的概率。

表 6-1 各节目主题分析结果

国家宝藏			中国诗词大会		
主题 1	主题 2	主题 3	主题 1	主题 2	主题 3
守护	泪目	陈晓	厉害	梦想	腹有诗书气自华
传承	难受	帅	武亦姝	光芒	喜欢
宝藏	就那样	喜欢	经典	孩子	感动
文化	融合	以为	背诵	人生	努力
以前	期待	乾隆	文化	读书	希望

朗读者			经典咏流传		
主题 1	主题 2	主题 3	主题 1	主题 2	主题 3
董卿	环球	遇见	毛晓彤	传唱	太白
胡歌	英语	人生	任嘉伦	演绎	苏轼
电视	中华	小伙伴	张杰	主题曲	中华
			王俊凯	好听	朱雀
			歌手	榜	传承

老师	传承	文艺
星光奖	正能量	安利

主题 1 中,《国家宝藏》反应了国家宝藏的核心在于守护宝藏、传承文化;《中国诗词大会》的观众觉得选手们是厉害的;《朗读者》的主持人以及嘉宾十分吸睛;《经典咏流传》的观众对节目演唱歌手张杰等的喜爱。

主题 2 中,《国家宝藏》反应了观众观看节目之后的感受;《中国诗词大会》反应大家认为这档节目非常适合孩子观看;《朗读者》反应观众希望节目朗读文章的范围可以放宽到全球;《经典咏流传》反应节目演绎的主题曲非常好听。

主题 3 中,《国家宝藏》反应了明星守护人和扮演的历史人物对观众有较大的吸引力;《中国诗词大会》反应了观众对这档节目的感受主要有喜欢、感动等;《朗读者》的观众觉得朗读者有浓厚的人文气息;《经典咏流传》观众开始关注歌曲背后的诗人及其他历史背景。

(四) 建议型评论分析

在收集数据的过程中,我们发现传统文化类综艺节目的新形式极大地激起了全民献策改进节目或者开拓节目新领域的热情,评论区内有大量的建议被提出。我们选取了这四个节目的官方微博(@CCTV 国家宝藏@CCTV 中国诗词大会@CCTV 朗读者@CCTV 经典咏流传)中置顶博文下的评论里点赞热度前三的建议型评论。(具体内容见附录一“热度前三的建议型评论”)。

七、结论与建议

(一) 研究结论

1.文化类综艺节目传播特征方面

(1)“公众人物+传统文化”模式助力传播

通过对新浪微博上传统文化类综艺节目的评论数据进行挖掘,我们发现在 13 万多的有效数据里有多达 58493 条都和若干明星嘉宾名字相关联,“公众人物效应”对节目起到了非常强的“吸粉”作用,由此,合适的嘉宾和核心价值观是制作一档优秀文化类节目的重中之重。

(2) 观众特征影响显著,高学历家庭观众观看意愿较高

二元回归模型结果得出,观众特征中观众年龄、教育程度等变量显著。接受过高等教育的群体是文化类综艺节目的主要观众,而在受教育程度低的群体中文

化类综艺节目观看率并不高，因此，如何降低文化类综艺节目的门槛，使更多的人参与进来成为节目制作是值得关注的问题。

2.文化类综艺节目传播效果方面

(1) 消除行业歧视，相关从业者更受尊重

观看节目后，认同该行业有前途的平均分从 3.35 提升至 5.01，居民对于传统文化从业者的待遇认知改变显著；认同该行业非主流的平均分从 4.02 下降至 2.71；节目让传统文化从业人员的负面印象降低。

(2) 传统文化领域受众年龄层改变，年轻人占比增加

从节目对观众行为影响的结果来看，观看节目后进行相关活动的人要远远多于观看前进行该活动的人（如日常讨论相关话题、查阅相关资料、游历文化古迹、参与相关行业或相关活动等，分别提升了 40.63%、56.25%、50%、26.56%）；从年龄层分布来看，传统文化领域受众年龄层年轻化，30 岁以下观众占比平均增加 24.8%。在观看传统文化类综艺节目之后，更多的年轻人开始关注相关领域。

(3) 节目特征显著，观众文化传承热爱及民族自豪感提高

Ridit 结果显示，使观众更有参与感的得分最高（0.5753），其次是明星入驻（0.5458），对文化传承有积极作用排名第三（0.5056）。这表明观众参与感和明星入驻受到观众的高度重视。同时，结构方程模型结果表明，节目内容专业性和观众参与程度等节目特征对观众文化传承的热爱和民族自豪感有显著的正效应。总体而言，文化类综艺节目的影响效果显著，对观众的文化传承热爱以及民族自豪感均有显著的促进作用。

(4) 节目成为新的文化教育形式

传统文化类综艺节目作为新型综艺模型在寓教于乐、传承传统文化方面显现出了极为可观的发展潜力。在认知渠道方面，67.64%的被调查观众选择了考纲要求，40.70%选择文化节目，说明文化类节目已经成为了除了考纲范围（如高考要求背诵的古诗文范围）和家庭教育以外的第三大认知渠道。

另外，自从此类节目播出之后，孩子对传统文化的兴趣增加，部分老师会在课上放一些文化类综艺节目的片段引起学生兴趣以及加强理解。由此可以看出，文化类综艺节目开拓了教学方式的新领域，并激起了学生得极大兴趣。

(5) 文物维护及推广经费增加，从业人员待遇提升

通过对博物馆从业者的深度访谈，工作人员表示自此类节目播出之后，来博物馆参观的游客明显增加，并且在节目中介绍过的藏品经常被参观，周边产品的销量增加，使得在展示场馆的日常维护等方面可投入资金增加。同时，有更多年轻人报名参加志愿者。对此我们可以看出，文化类综艺节目对社会具有正的影响。

另外，文化类综艺节目播出后掀起了一股传统文化热，由于民众的需求增加，国家在此方面也增加了文化产业方面资金的供给，其中相当一部分比例用于文物

的维护和推广。根据国家统计局年度数据，2016年，2017年国家财政文化支出分别增长2.81%和6.18%，

（7）观众二次传播意愿较高，掀起全民献策热潮

82.81%的观众在观看节目后愿意继续关注此类节目，用户粘性极高，利于节目的继续发展；84.38%的人愿意讲节目推荐给他人，二次传播效果极好，利于节目在群众中的自行普及传播。

在数据挖掘对数据进行分析的过程中，我们发现传统文化类综艺节目的新形式极大地激起了全民献策改进节目或者开拓节目新领域的热情，评论区内有大量的建议被提出。

（二）建议

1.创新节目形式，避免审美疲劳

在不会持续关注节目的人群中有72.73%的人认为节目同质化严重，因此节目应该采取差异化定位，避免同质化，适当提高节目的原创性，节目组需要不断创新，对节目的形式和内容上不断的做出调整，譬如增加与观众的互动方式、采用恰当的娱乐表现形式等，以此避免观众审美疲劳，让文化类综艺节目在市场上得以长远的发展。

2.挖掘新的传统文化领域

文化类综艺节目作为市场上“传统文化”+“综艺”的一种创新，这类节目应该始终秉持着做好文化内涵这一核心，挖掘优秀传统文化，将节目的形式不仅仅局限于唱歌、读诗，正如问卷的结果所示，73%观众认为可以增加历史文化名城名镇名村等的解读，67%的观众认为可以对中华老字号做多一点的挖掘，58%的观众觉得有必要增加关于节庆礼仪服装服饰的节目，另外还有一部分观众认为节目形式上进行创新，譬如增加室外的节目。这些都是文化类综艺节目的创新方式，提升节目内涵，综合利用多种文化元素。但是，在创新的同时，要始终秉持着节目需要承担社会责任、传播正能量的信念，不断的为观众挖掘优秀传统文化。

3.借力公众人物效应捆绑传播传统文化

问卷调查的结果显示在观看节目之后，仍有部分的人不会被节目产生任何行为的影响效果，例如：有26.56%的观众不会参与传统文化相关行业或相关活动。我们建议借力公众人物效应捆绑传播传统文化，在文化类综艺节目刚开始播出和宣传建立口碑的时候，主打“公众人物+传统文化”的模式，吸引观众，再通过明星的身体力行，以及传统文化自身的魅力，引导大众在生活中参与文化相关活动，增强节目传播效果。

4.利用新媒体平台增强观众对节目的参与感

在节目传播渠道的调查中我们发现57.69%的被调查者，选择了微博、朋友圈等新媒体为传统文化类综艺节目的认知渠道；同时，通过Ridit分析法，节

目参与性是观众最为认可的节目闪光点。结合二者，我们建议节目通过新媒体平台增加节目的参与感，如：建立官方微博、官方微信公众号，利用其传播节目播出时间等信息的同时，增设如小游戏、互动抽奖、接纳观众投稿等互动性强的活动；增加与微博、微信公众号中有影响力的平台的合作，扩大曝光度的同时，让节目的线上活动有更多人参与，提升观众参与感。

5.降低文化壁垒，加强文化输出

通过对外国观众的访谈，我们发现阻碍文化类综艺节目向海外大力度传播的原因是因为外国观众因为语言原因在节目内容的理解上存在一定的困难，以及存在知识和文化壁垒。Youtube 上的外国观众看中国文化类综艺视频评论里时常出现博主们非常感兴趣非常想看但又心有余而力不足的情况。因此我们建议节目增加字幕翻译，或稍微降低节目的理解难度，增加新的海外版本，帮助外国观众理解节目的内容，以此增强我国的文化输出和文化软实力。

6.开拓播放渠道，增强宣传力度

通过观众籍贯对电视台偏好的影响，我们发现某个地区的人对本地区的电视台有偏好，并且籍贯为北方的观众相对于南方的观众更喜爱看中央电视台。我们建议地方电视台多引进文化类综艺节目，并且在某个固定的时间段播放，让更多的居民接触这类节目。同时，节目可以更加注重宣传，打造更高的知名度。

参考文献

- [1] 刘娅蒙.文化类综艺节目研究现状梳理[J].西部广播电视,2017,(19)
- [2] 刘兴林.基于定位理论研究我国文化类节目的发展策略[J].新闻传播,2017,(7)
- [3] 蒋婷婷.文化类电视综艺节目的创意与传播研究[D].曲阜师范大学,2016
- [4] 刘冰,王新雨.文化类综艺节目收受调查[J].青年记者,2017,(19)
- [5] 孙晴.文化类综艺节目的困境与破局[J].北方文学（下旬刊）,2017,(3)
- [6] 郭学暖.从《中国面孔》看文化类综艺节目的创新[J].青年记者,2015,(2)
- [7] 李金昌.应用抽样技术.科学出版社,2015.1
- [8] Willson, V. L. Research technique sin AER Jarticles. Educational Researcher, 9(6)
- [9] Aiken, L. R. Psychological testing and assessment (9th ed.)
- [10] 张琳,郭玉真.文化类综艺节目的文化价值[J].青年记者,2017,(23)
- [11] 刘林.浅谈文化类综艺节目荧屏热播现象[J].活力,2017,(6)
- [12] 杨丽娟.文化类综艺节目的传播特色[J].声屏世界,2015,(9)
- [13] 侯泽武.从《国家宝藏》来看文化类综艺节目的创新[J].北方文学（刊）,2018,(1)

- [14]张丹.文化类综艺节目的爆发[J].新闻研究导刊,2017,(10)
- [15]张丹.文化类综艺节目的创新与发展[J].新闻传播,2017,(18)
- [16]孟笛,李晓倩.从《朗读者》看我国文化类综艺节目发展现状[J].西部广播电视,2017,(20)
- [17]胡奇军.《国家宝藏》:传统文化的现代性表达[J].电影评介,2018,(1)
- [18]郭增强.审视大众媒介的社会责任——以央视文化类综艺节目为例[J].魅力中国,2017,(z1)
- [19]周璞.浅析文化类综艺节目的成功之道——以《见字如面》为例[J].听,2017,(5)
- [20]张丹.文化类综艺节目的爆发[J].新闻研究导刊,2017,(10)
- [21]胡函博.文化类综艺节目的文化传承与艺术呈现——以《中国诗词大会》为例[J].辽宁经济,2017,(10)
- [22]别君华.文化类节目的创新传播策略——以《朗读者》为例[J].传媒,2017,(15)
- [23]金妍.坚守与创新:传统文化节目的升级路径[J].青年记者,2017,(17)
- [24]王崑欣.以《朗读者》为例探析文化类综艺节目的发展定位[J].西部广播电视,2017,(8)
- [25]常诚.浅析文化类电视节目对表演与文化元素的创新融合运用——以《见字如面》为例[J].传媒,2017,(12)
- [26]彭品荣.大众审美文化视域下"梦想类"综艺节目的"梦想"解读[J].剑南文学,2015,(8)
- [27]宋彦丽.由单一到多面:文化类电视节目的转型升级[J].视听,2018,(2)
- [28]陈瑞.试析文化综艺节目的品质跃升与艺术创新策略[J].浙江传媒学院学报,2017,(3)
- [29]曾筱霄.试论坚定文化自信的路径——以文化综艺类节目为视角[J].开封教育学院学报,2017,(6)
- [30]过彤,张庆龙.《朗读者》:文化类电视综艺节目的大众化探索[J].传媒评论,2017,(3)
- [31]戴硕.读书类电视节目的形态创新策略研究——以《朗读者》为例[J].创作与评论,2017,(10)
- [32]文卫华,王晶晶.原创文化节目的现实图景与发展路径[J].青年记者,2017,(9)
- [33]肖芄.朗读类节目:从文字到影像的意义增值[J].现代传播,2017,(8)
- [34]王燕,刘悦,欧阳瀚.传统文化在综艺节目中的传播研究[J].西部广播电

视,2017,(21)

[35]解晶晶.浅谈电视综艺节目的发展趋势[J].商情,2017,(42)

附录一

（一）调查方式与方法

1.调查方式

在 $n > 2$ 时，严格的 πPS 抽样法^[7]计算难度较大，实施成本较高，故我们在本次调查中采取多阶段 $n = 5$ 的非严格的 πPS 抽样法，即固定样本 $n = 5$ ，允许一阶包含概率 π_j 与总体呈近似比例的，每个单位在每次抽样中入样的概率与单位的规模大小成比例的不放回抽样。

在此次 πPS 抽样中，我们采取耶茨-格伦迪法^[7]进行抽取：

第一个样本单位按 Z_i 的概率抽取，设第 i 个单位入样；

第二个样本单位按 $\frac{Z_j}{1-Z_i}$ 的概率在余下 $N-1$ 个单位中抽取，设第 j 个单位入样；

第三个样本单位按 $Z_k / (1-Z_i - Z_j)$ 的概率在剩下的 $N-2$ 个单位中抽样；

以此类推。

由于武汉市各区的居民对于文化类节目的观看情况没有特别大的差异，故以直接根据武汉市的行政分区进行编号。又考虑到各区人口差距较大，故以人口为辅助变量。

第一阶段，在武汉市中，以各行政区划分，采用 $n = 5$ 的非严格的 πPS 抽样法，抽取出入样行政区；

第二阶段，在被抽中的行政区中，以各街道划分，抽取出入样街道；

第三阶段，在被抽中的街道中，以各社区划分，抽取出入样社区；

第四阶段，对于第三阶段抽中的街道，前往相应居委会。通过居委会获得住户信息并编制抽样框。采用随机起点的等距抽样中的直线抽取法，通过最优样本量以及该街道总人口确定抽样间距，再由计算机产生间距内一随机数，从而确定入样样本。

针对入户时户主不在家或是拒绝接受访问的用户，我们将以入样住户最临近的右侧住户入样。

2.调查方法

本次调查中，我们根据研究的需求，采取了文案调查法、深度访谈法、文本挖掘法进行定性研究；采用问卷形式的面访进行定量研究。

（二）文本挖掘

1.数据清洗与分词

在数据挖掘中，海量的原始数据中存在着大量不完整、不一致、有异常的数据，严重影响到数据挖掘建模的执行效率，甚至可能导致挖掘结果的偏差。具体对于微博文本数据来说，由于某些节目请的嘉宾明星粉丝较为狂热，为自己的偶像疯狂发布含有节目关键词的博文，导致关键词搜索含有大量此类无关文本，因此进行原始数据的清洗尤为重要。

本次预处理的主要工作是清除粉丝军团发表的带有节目关键词但和节目本身无关的数据处理规则如下：

- ① 评论文本内容与用户名均相同的数据视为粉丝刷屏发表的，予以删除。

微博内容	微博发布者	搜索关键字
#胡歌##胡歌 朗读者 ##最美好的胡歌#@胡歌@胡歌@胡歌我在#明星 ALL 榜#上为你加油啦，你是我今生唯一的执著哦。#胡歌#棒棒哒！@超级应援网页链接	小倩 nought	朗读者
#胡歌##胡歌 朗读者 ##最美好的胡歌#@胡歌@胡歌@胡歌我在#明星 ALL 榜#上为你加油啦，你是我今生唯一的执著哦。#胡歌#棒棒哒！@超级应援网页链接	小倩 nought	朗读者
#胡歌##胡歌 朗读者 ##最美好的胡歌#@胡歌@胡歌@胡歌我在#明星 ALL 榜#上为你加油啦，你是我今生唯一的执著哦。#胡歌#棒棒哒！@超级应援网页链接	小倩 nought	朗读者

- ② 评论内容完全和节目无关，仅有节目关键词，视为粉丝随意发表的评论，予以删除。

微博内容	微博发布者	搜索关键字
#胡歌 朗读者 ##胡歌如梦之梦##胡歌猎场#众星捧月星光熠熠生日会小胡歌大理想	白玉兰视帝胡歌	朗读者
#胡歌##胡歌猎场##胡歌 朗读者 #我们还有个朗读者的粮呢？我们还有个上海纪录片呢？？我们在啃大猎场那是因为小祖宗不出来如果如果拍戏表示我也很想体验一把官宣的感觉@胡歌胡歌微博美...	破晓潇雨	朗读者

经过预处理后，共得到评论数据 134915 条。无效的数据达 123493 条，文本有效率仅为 52.21%。

不同于在英文中，每个单词之间都有空格来隔开。在中文里，只有字、句和段落能够通过明显的分界符进行快速的划界，对于“词”和“词组”来说，其边界模糊没有一个形式上的分界符。而词是最小的，能够独立运用的有意义的语言单位，文本分词的工作也就是将连续的字序列按照一定的规范重新组合成词序列。本文的分词工作是使用 R 软件的 Rwordseg 包来完成的，基本实现原理及步骤如下：

①一二节清洗：将文本内容转换成向量之后，将其中包含的数字、英文字符、全英文的 dot 符号、空值文本、字符数小于 2 的文本都循环去除。

②分词：导入进行过一二节清洗的文本数据之后，使用 Rwordseg 包中的的 segmentCN 函数进行分词处理，为了避免一些关键词无法识别导致被切割成无意义单字，我们从搜狗中文网上面下载了明星名字和国家宝藏以及综艺相关的词典，加入词库后再进行分词。

③清洗停用词：分词后的词语中含有大量的“了”、“的”、“这”等高频无意义词语，为了保证分析的科学性，我们将这些停用词予以删除。

（相关 R 代码见附录五）

2. 热度前三的建议型评论

表 6-2 中国诗词大会评论建议

想说一下宣传的问题。在电视上有宣传。在网上，因为网上比如微博，宣传费高的惊人，节目组可能想把钱用在舞台、内容等方面。我觉得这样也好，把买热搜的钱用在选手的食宿和评委的酬劳，以及工作人员的奖金上，也都是很好的。	榆树钱儿	6564
诗词大会下一季建议放到暑期播放，或许更合适！	bill_bill	2938

表 7-3 朗读者评论建议

微博内容	微博发布者	点赞数
作为一个语文老师表示朗读者已经成为了我学生们必备的作文素材库，希望朗读者的朗读范围可以再拓宽一点，不要太局限了	心也 Hoo	6546
什么时候有英文版啊什么时候有英文朗读稿啊！！	磨人的傻兔子	5648
希望朗读者可以请更多有文化气息和功底的明星来演绎，流量小生是真的不行，完全无法入戏	盼小盼 Q	1922

表 7-4 经典咏流传评论建议

微博内容	微博发布者	点赞数
我觉得点红心这个环节真的有点多余。尤其是点评嘉宾点红心搞得像选秀比赛一样。如果是想表现与观众的互动可以想想更好的办法，再下点功夫。	靓仔嘅虎牙	3540
都看完了，也非常喜欢这个节目。但说句不太招人待见的话，如果是真正的咏唱古诗词，比如《青玉案》这种，而不是夹杂着现代歌词，我会更喜欢，而且可以直接给孩子们听，会唱就会背了 23333	Archimedes	2501
央视这样宣传诗词为什么不开一档讲南派吟诵的节目呢？古诗难道不是应该用吟的方式吟出来么	落叶扁舟	935

附录二

表 2-2 武汉市居民入户调查抽样框

第一阶段 抽样框	第一阶段 入样单位	第二阶段 抽样框	第二阶段 入样单位	第三阶段抽 样框	第三阶段 入样单位	第四阶段 抽样框
武汉市 13个行政 区	江汉区	江汉区 所有 14个街道	积玉桥街道	积玉桥街道 16个社区	汉成里社区 凤凰社区 沙湖村社区	汉成里社区所有居民 凤凰社区所有居民 沙湖村社区所有居民
			中南路街道	中南路街道 31个社区	小龟山社区 静安社区 涂家岭社区	小龟山社区所有居民 静安社区所有居民 涂家岭社区所有居民
			珞珈山街道	珞珈山街道 6个社区	东湖村社区 珈智社区 茶港社区	东湖村社区所有居民 珈智社区所有居民 茶港社区所有居民
	江夏区	江夏区 所有 12个街道	金口街道	金口街道 9个社区	花园社区 文字港社区 杨园社区	花园社区居民 文字港社区居民 杨园社区居民
			乌龙泉街道	乌龙泉街道 26个社区	五星村 洪湾村 乌龙泉社区	五星村所有居民 洪湾村所有居民 乌龙泉社区所有居民
			纸坊街道	纸坊街道 17个社区	龙井社区 幸福居委会 宁安社区	龙井社区所有居民 幸福社区所有居民 宁安社区所有居民
	洪山区	洪山区 所有 8个街道	关山街道	关山街道 38个社区	磨山社区 关山路社区 鲁广社区	磨山社区所有居民 关山路社区所有居民 鲁广社区所有居民
			卓刀泉街道	卓刀泉街道 19个社区	体院社区 金泉社区 吴家湾社区	体院社区所有居民 金泉社区所有居民 吴家湾社区所有居民
			狮子山街道	狮子山街道 11个社区	华农东社区 珞狮路社区 玫瑰湾社区	华农东社区所有居民 珞狮路社区所有居民 玫瑰湾社区所有居民
	汉阳区	汉阳区 所有 11个街道	晴川街道	晴川街道 5个社区	国棉社区 江汉桥社区 洗马社区	国棉社区所有居民 江汉桥社区所有居民 洗马社区所有居民
			永丰街道	永丰街道 18个社区	玉乐社区 玉龙苑社区 仓吉社区	玉乐社区所有居民 玉龙苑社区所有居民 仓吉社区所有居民
			江堤街道	江堤街道 13个社区	向阳村社区 丰收社区 新新社区	向阳村社区所有居民 丰收社区所有居民 新新社区所有居民
	江岸区	江岸区 所有 16个街道	大智街道	大智街道 7个社区	宁波社区 先锋社区 韩家社区	宁波社区所有居民 先锋社区所有居民 韩家社区所有居民
			球场街道	球场街道 5个社区	安静社区 球新社区 红艳社区	安静社区所有居民 球新社区所有居民 红艳社区所有居民
			劳动街道	劳动街道 10个社区	三合社区 解放社区	三合社区所有居民 解放社区所有居民

表 3-1 调查实施进度表

起止时间	调查步骤与内容	参与人员
2017.12.18—2018.3.3	搜集热点话题和相关资料,小组成员讨论后确定选题;确定研究思路和调查目的,并据其制定研究假设与调查内容;搜集文献与资料,了解文化类综艺节目的现状。	队长负责 全体成员
3.4—3.7	设计调查方案和调查问卷,咨询指导老师修改、确认方案与问卷;选择抽样方案,确定抽样框。	成员四
3.7	根据入样的城区和街道,考虑时间和成本因素,制定具体调查日程安排。	成员三
3.8	打印调查问卷,购买圆珠笔、文件夹等文具	成员二
3.9	确定质量控制措施,对本组成员进行相关技能培训,分配每个人的调查任务。	成员五
3.10—3.24	进入正式调查阶段,通过入户访问对调查单位发放问卷,获取有效信息。	全体成员
3.25—3.31	利用R等统计软件,进行相关数据处理与建模分析。	成员三、五
4.1—4.9	撰写调研报告,审核并修改。	全体成员

表 4-3 二元选择模型变量

变量名	变量符号	变量定义
最近一年内是否观看过文化类综艺节目	Y	$Y = \begin{cases} 1, & \text{观看过} \\ 0, & \text{从未观看} \end{cases}$
性别	Male	$Male = \begin{cases} 1, & \text{男} \\ 0, & \text{女} \end{cases}$
年龄	Age20_29 Age30_39	$Age20_29 = \begin{cases} 1, & 20-29\text{岁} \\ 0, & \text{其他} \end{cases}$

	续表 4-3 二元选择模型变量	$Age_{30_39} = \begin{cases} 1, & 30-39\text{岁} \\ 0, & \text{其他} \end{cases}$
变量名	变量符号	变量定义
最高学历	High	$High = \begin{cases} 1, & \text{高中/中专/职校} \\ 0, & \text{其他} \end{cases}$
	College	$College = \begin{cases} 1, & \text{大学专科} \\ 0, & \text{其他} \end{cases}$
	Undergraduate	$Undergraduate = \begin{cases} 1, & \text{大学本科} \\ 0, & \text{其他} \end{cases}$
	Postgraduate	$Postgraduate = \begin{cases} 1, & \text{硕士及以上} \\ 0, & \text{其他} \end{cases}$ 初中及以下为基础类型
职业	Job1	$Job1 = \begin{cases} 1, & \text{普通职员} \\ 0, & \text{其他} \end{cases}$
	Job2	$Job2 = \begin{cases} 1, & \text{企业管理员} \\ 0, & \text{其他} \end{cases}$
	Job3	$Job3 = \begin{cases} 1, & \text{政府/机关干部/公务员} \\ 0, & \text{其他} \end{cases}$
	Job4	$Job4 = \begin{cases} 1, & \text{自由职业者} \\ 0, & \text{其他} \end{cases}$
	Job5	$Job5 = \begin{cases} 1, & \text{个体经营者/承包商} \\ 0, & \text{其他} \end{cases}$
职业	Job6	$Job6 = \begin{cases} 1, & \text{普通工人} \\ 0, & \text{其他} \end{cases}$

变量名	变量符号	变量定义
职业	Job7 Job8 续表 4-3 二元选择模型变量	$Job7 = \begin{cases} 1, & \text{商业服务业职工} \\ 0, & \text{其他} \end{cases}$ $Job8 = \begin{cases} 1, & \text{农林牧渔劳动者} \\ 0, & \text{其他} \end{cases}$ 全日制学生为基础类型
家中未成年孩子数量	Kid1 Kid2 Kid3	$Kid1 = \begin{cases} 1, & \text{1个孩子} \\ 0, & \text{其他} \end{cases}$ $Kid2 = \begin{cases} 1, & \text{2个孩子} \\ 0, & \text{其他} \end{cases}$ $Kid3 = \begin{cases} 1, & \text{3个孩子} \\ 0, & \text{其他} \end{cases}$ 无为基础类型
对传统文化认知的来源	Source1 Source2 Source3 Source4	$Source1 = \begin{cases} 1, & \text{家庭教育} \\ 0, & \text{其他} \end{cases}$ $Source2 = \begin{cases} 1, & \text{教育} \\ 0, & \text{其他} \end{cases}$ $Source3 = \begin{cases} 1, & \text{自身兴趣} \\ 0, & \text{其他} \end{cases}$ $Source4 = \begin{cases} 1, & \text{文化节目} \\ 0, & \text{其他} \end{cases}$ 考纲要求为基础类型
籍贯	South	$South = \begin{cases} 1, & \text{华南地区} \\ 0, & \text{其他} \end{cases}$

续表 4-3 二元选择模型变量		
变量名	变量符号	变量定义
籍贯	Middle	$Middle = \begin{cases} 1, & \text{华中地区} \\ 0, & \text{其他} \end{cases}$
	North	$North = \begin{cases} 1, & \text{华北地区} \\ 0, & \text{其他} \end{cases}$
	Northwest	$Northwest = \begin{cases} 1, & \text{西北地区} \\ 0, & \text{其他} \end{cases}$
	Southwest	$Southwest = \begin{cases} 1, & \text{西南地区} \\ 0, & \text{其他} \end{cases}$
	Northeast	$Northeast = \begin{cases} 1, & \text{东北地区} \\ 0, & \text{其他} \end{cases}$
华东地区为基础类型		

表 3-2 各题项高低组占比

题项	低分组占比 (%)	高分组占比 (%)
Q17.1 趣味性	25.64	74.36
Q17.2 参与感	39.74	60.26
Q17.3 明星入驻	30.77	69.23
Q17.4 内容专业性	23.08	76.92
Q17.5 传承积极作用	33.33	66.67

表 3-3 Cronbach α 信度系数预调查检验

层面	Cronbach α 系数	项数	信度评价
总体	0.831	5	信度好

表 3-4 折半信度系数预调查检验

Cronbach α 系数	第一部分	值	0.790
		项数	3a
	第二部分	值	0.621
		项数	2b
	总项数		5
形态之间的相关性			0.661
斯皮尔曼-布朗系数	等长		0.796
	不等长		0.802
格特曼折半系数			0.763

表 3-5 单项与总和和相关效度预调查分析

	趣味性	参与感	明星入驻	内容专业性	传承积极作用
相关系数	0.478	0.242	0.327	0.336	0.405
P 值(双侧)	0.000	0.032	0.004	0.003	0.000

表 3-6 预调查 KMO 和 Bartlett 球形度检验

KMO 取样適切性量数		.602
巴特利特球形度检验	近似卡方	86.227
	自由度	28
	显著性	.000

表 3-7 观众眼中的节目闪光点与年龄阶段列联表

	19 岁及 以下	20-29	30-39	40-49	50-59	60 岁及 以上
较传统类节目更具 趣味性	27	124	152	72	42	21

观众更有参与感	15	231	125	56	50	25
明星入驻	36	217	145	45	37	24
内容专业性	14	157	169	74	67	57
对文化传承有积极作用	19	258	201	82	78	64
其他	5	43	61	13	15	19

表 4-4 Logit 模型对是否观看节目的回归显著结果

变量	系数	标准差	Wald 统计量	P 值
Male	0.0599	0.1159	0.52	0.608
Age20_29	-0.0860	0.1188	-0.72	0.473
Age30_39	-0.0773	0.1299	-0.59	0.555
Age40_49	-0.6303	0.2416	-2.61	0.012**
Age50_59	-0.1082	0.1376	-0.79	0.436
Age60	0.6264	0.4101	1.53	0.134
High	-0.2355	0.2097	-1.12	0.267
College	-0.0547	0.2134	-0.26	0.799
Undergraduate	-0.1132	0.1625	-0.7	0.489
Postgraduate	0.7076	0.2293	3.09	0.003***
Job2	0.3330	0.2242	1.49	0.144
Job3	0.2275	0.3044	0.75	0.459
Job4	0.4433	0.2830	1.57	0.124
Job5	0.0178	0.2849	0.06	0.951
Job6	0.0439	0.3071	0.14	0.887
Job7	0.3778	0.3578	1.06	0.296
Job8	-0.0003	0.1883	0	0.999
Job9	0.3995	0.3359	1.19	0.241
Kid1	0.4530	0.2525	1.79	0.079*

续表 4-4 logit 模型对是否观看节目的回归结果

Kid2	0.0636	0.1131	0.56	0.577
Kid3	0.0893	0.1517	0.59	0.559
Source1	-0.3714	0.3230	-1.15	0.254
Source2	-0.0214	0.1363	-0.16	0.876
Source3	0.1674	0.1365	1.23	0.227
Source4	0.1217	0.1642	0.74	0.462
South	0.0045	0.1823	0.02	0.98
变量	系数	标准差	Wald 统计量	P 值
Middle	0.0055	0.1603	0.03	0.973
North	-0.0803	0.1937	-0.41	0.68
Northwest	-0.0095	0.1584	-0.06	0.952
Southwest	0.1075	0.2879	0.37	0.71
Northeast	0.2599	0.2251	1.15	0.254
Cons	0.9138	0.2460	3.71	0.001***
R2=0.4260				

表 5-1 轮廓分析结果表

效应	值	F	假设df	误差df	Sig.
Pillai的跟踪	.950	89.728 ^b	4.000	19.000	.000
Wilks的Lambda	.050	89.728 ^b	4.000	19.000	.000
因子*g Hotelling 的跟踪	18.890	89.728 ^b	4.000	19.000	.000
Roy的最大根	18.890	89.728 ^b	4.000	19.000	.000

附录三

“钟灵毓秀 曲水流觞”文化类综艺节目传播特征与传播效果调查问卷

问卷编号： 调研员： 时间： 调研地点：

尊敬的朋友：

您好！我们是在校大学生，为了参加“第八届全国大学生市场调查与分析大赛”，特以“文化类综艺节目”为主题展开调查活动。本次调查将对您提供的信息进行严格保密，不用写调查者姓名，所有回答用于学术研究。您只需根据自己的实际情况填写本问卷。衷心感谢您的支持与配合！

注：无特别说明，请直接在选中答案前的“□”内打√。

◎第一部分：基础数据

您的性别：男 女

您的年龄：19岁及以下 20-29岁 30-39岁 40-49岁
50-59岁 60岁及60岁以上

最高学历：初中及以下 高中/中专/职校 大学专科 大学本科
硕士及以上

您的职业：在校学生 普通职员 企业管理员 政府/机关干部/公务员
自由职业者 个体经营者/承包商 普通工人
商业服务业职工 农林牧渔劳动者 退休
其他：_____

您家有多少个孩子（未成年）：无 1个 2个 3个及以上

您的传统文化认知的来源（多选）：

考纲要求 家庭教育 旅游 自身兴趣 文化节目 其他

您平常喜欢观看哪些电视台？（多选）

湖南卫视 浙江卫视 东方卫视 江苏卫视 北京卫视
辽宁卫视 山东卫视 河北卫视 河南卫视 安徽卫视
江西卫视 湖北卫视 陕西卫视 天津卫视 云南卫视
四川卫视 重庆卫视 中央电视台__套（1-16，可填多个）
其他_____

您的籍贯：_____（省份地区）

◎第二部分：文化类综艺节目观看情况及偏好

1、请问您最近一年内是否观看过文化类综艺节目？

是 否

2、您喜欢看下列文化类综艺节目中的哪些？（多选）

- 中国诗词大会 国家宝藏 朗读者 见字如面 儿行千里
中国戏曲大会 中国汉字听写大会 汉字风云会 汉语桥
非凡匠心 如果国宝会说话 向上吧诗词 诗书中华
世界听我说 念念不忘 小镇故事 文明之旅
绿水青山看中国 群英会 世界青年说 其他_____

3、您是否会鼓励家里的小孩看这一类节目？

是 否

4、请问您是如何知道文化类综艺节目的？（多选）

- 微博/微信朋友圈/公众号推送 新闻/报纸杂志
朋友/亲人/同学告知 电视广告宣传
明星宣传 无意中得知 其他_____

5、以下为您在观看文化类综艺节目时可能会考虑的因素，其重要程度是：（请在相应的等级下打勾）

	非常不重要 ☆	不重要 ☆☆	普通 ☆☆☆	重要 ☆☆☆☆	非常重要 ☆☆☆☆☆
电视台影响力	<input type="checkbox"/>				
节目形式	<input type="checkbox"/>				
内容质量	<input type="checkbox"/>				
舞台效果	<input type="checkbox"/>				
播出时间段	<input type="checkbox"/>				

6、您观看此类节目的频率：

从未观看 一周小于1小时 一周1-3小时 一周3小时以上

7、您对此类节目的看法：

仅供娱乐 传承文化底蕴 寓教于乐 其他

观看过文化类综艺节目的观众填写 8-20 题

8、您认为此类节目的闪光点进行打分：

	非常不重要 ☆	不重要 ☆☆	普通 ☆☆☆	重要 ☆☆☆☆	非常重要 ☆☆☆☆☆
趣味性	<input type="checkbox"/>				
参与感	<input type="checkbox"/>				
明星入驻	<input type="checkbox"/>				
内容专业性	<input type="checkbox"/>				
传承积极作用	<input type="checkbox"/>				
电视台影响力	<input type="checkbox"/>				
舞台效果	<input type="checkbox"/>				
播出时间段	<input type="checkbox"/>				

9、您会在日常生活中讨论与文化类综艺节目相关的话题？

以前就会 否 是

10、您是否会查阅相关资料？

以前就会 否 是

11、您是否会游历文化古迹（参观博物馆或名人故居等）？

以前就会 否 是

12、您是否会参与传统文化相关行业或相关活动（例如志愿者）？

以前就会 否 是

13、观看此类节目对您的民族自豪感造成了什么影响？（量表题）

降低很多	稍有降低	没有影响	稍有提升	提升很大
-2	-1	0	1	2

14、观看此类节目使你对传统文化的热爱产生了什么影响？

降低很多	稍有降低	没有影响	稍有提升	提升很大
-2	-1	0	1	2

15、请判断下面的每个形容词是否符合你在观看节目之前对传统文化相关类行业从业者（如文物修复师、手艺人等）的看法。（★1—7 非常不符合—

—非常符合)

非主流	☆☆☆☆☆☆☆
有前途	☆☆☆☆☆☆☆
低收入	☆☆☆☆☆☆☆
温文尔雅	☆☆☆☆☆☆☆
老练得体	☆☆☆☆☆☆☆
雄心勃勃	☆☆☆☆☆☆☆
无能的	☆☆☆☆☆☆☆
像个领导的	☆☆☆☆☆☆☆
腐朽	☆☆☆☆☆☆☆
铁饭碗	☆☆☆☆☆☆☆
稳定	☆☆☆☆☆☆☆
福利好待遇高	☆☆☆☆☆☆☆
知识渊博	☆☆☆☆☆☆☆
无私	☆☆☆☆☆☆☆
多才多艺	☆☆☆☆☆☆☆
辛苦劳累	☆☆☆☆☆☆☆

16、请判断下面的每个形容词是否符合你在观看节目之后对传统文化相关类行业从业者（如文物修复师、手艺传承人等）的看法。（★1—7 非常不符合—

—非常符合)

非主流	☆☆☆☆☆☆☆
有前途	☆☆☆☆☆☆☆
低收入	☆☆☆☆☆☆☆
温文尔雅	☆☆☆☆☆☆☆
老练得体	☆☆☆☆☆☆☆
雄心勃勃	☆☆☆☆☆☆☆
无能的	☆☆☆☆☆☆☆
像个领导的	☆☆☆☆☆☆☆
腐朽	☆☆☆☆☆☆☆

铁饭碗	☆☆☆☆☆☆☆
稳定	☆☆☆☆☆☆☆
福利好待遇高	☆☆☆☆☆☆☆
知识渊博	☆☆☆☆☆☆☆
无私	☆☆☆☆☆☆☆
多才多艺	☆☆☆☆☆☆☆
辛苦劳累	☆☆☆☆☆☆☆

17、您是否会持续关注此类节目？

是 否

18、您不会持续关注的原因是（会持续关注的观众不用填）（多选）：

审美疲劳 节目同质化严重 内容平庸 节目不专业
节目情感空泛 其他

19、你是否会将此类节目推荐给他人看？

是 否

未观看过文化类综艺节目的观众填写第 20 题

20、您希望此类节目增加以下哪些主题领域（多选）：

历史文化名城、名镇、名村
方言文化
中华老字号
节庆礼仪服装服饰
历史建筑
民族传统体育项目
其他

我们的问卷到此结束，衷心感谢您的参与！

附录四

问卷编码表

变量序号	变量名	变量类型	取值范围	选项	对应问题	对应题号
1	Q1	数值型	0-2	1-男, 2-女, 0-缺失	性别	
2	Q2	数值型	0-6	1-19岁及以下, 2-20-29岁, 3-30-39岁, 3-40-49岁, 4-50-59岁, 5-60岁及以上, 0-缺失	年龄	
3	Q3	数值型	0-5	1-初中及以下, 2-高中/中专/职校, 3-大学专科, 4-大学本科, 5-硕士及以上, 0-缺失	最高学历	
4	Q4	数值型	0-11	1-在校学生, 2-普通职员, 3-企业管理人员, 4-政府/机关干部/公务员, 5-自由职业者, 6-个体经营者/承包商, 7-普通工人, 8-商业服务业职工, 9-农林牧渔劳动者, 10-退休, 11-其他, 0-缺失	职业	
5	Q5	数值型	0-4	1-无, 2-1个, 3-2个, 3-3个及以上, 0-缺失	未成年孩子数量	
6	Q6	数值型	0-6	1-考纲要求, 2-家庭教育, 3-旅游, 4-自身兴趣, 5-文化节目, 6-其他, 0-缺失	传统文化认知的来源	
7	Q7	数值型	0-19	1-湖南卫视, 2-浙江卫视, 3-东方卫视, 4-江苏卫视, 5-北京卫视, 6-辽宁卫视, 7-山东卫视, 8-河北卫视, 9-河南卫视, 10-安徽卫视, 11-江西卫视, 12-湖北卫视, 13-陕西卫视, 14-天津卫视, 15-云南卫视, 16-四	平时最爱观看的电视台	

续编码表:

				川卫视, 17-重庆卫视, 18-中央电视台, 19-其他, 0-缺失		
8	Q9	文字值			籍贯	
9	Q10	数值型	0-2	1-是, 2-否, 0-缺失	您最近一年内是否观看过文化类综艺节目	1
10	Q11	数值型	0-21	1-中国诗词大会, 2-国家宝藏, 3-朗读者, 4-见字如面, 5-儿行千里, 6-中国戏曲大会, 7-中国汉字听写大会, 8-汉字风云会, 9-汉语桥, 10-非凡匠心, 11-如果国宝会说话, 12-向上吧诗词, 13-诗书中华, 14-世界听我说, 15-念念不忘, 16-小镇故事, 17-文明之旅, 18-绿水青山看中国, 19-群英会, 20-世界青年说, 21-其他, 0-其他	您喜欢看下列文化类综艺节目中的哪些	2
11	Q12	数值型	0-2	1-是, 2-否, 0-缺失	您是否会鼓励家里的小孩看这一类节目	3
12	Q13	数值型	0-7	1-微博/微信朋友圈/公众号推送, 2-新闻/报纸杂志, 3-朋友/亲人/同学告知, 4-电视广告宣传, 5-明星宣传, 6-无意中得知, 7-其他, 0-缺失	您是如何知道文化类综艺节目的	4
13	Q14.1	数值	0-5	1-非常不重要, 2-不重要,	观看文化	5

续编码表:

		型		3-普通, 4-重要, 5-非常重要, 0-缺失	类综艺节目时, 认为电视台影响力的重要程度	
14	Q14.2	数值型	0-5	1-非常不重要, 2-不重要, 3-普通, 4-重要, 5-非常重要, 0-缺失	观看文化类综艺节目时, 认为节目形式的重要程度	5
15	Q14.3	数值型	0-5	1-非常不重要, 2-不重要, 3-普通, 4-重要, 5-非常重要, 0-缺失	观看文化类综艺节目时, 认为内容质量的重要程度	5
16	Q14.4	数值型	0-5	1-非常不重要, 2-不重要, 3-普通, 4-重要, 5-非常重要, 0-缺失	观看文化类综艺节目时, 认为舞台效果的重要程度	5
17	Q14.5	数值型	0-5	1-非常不重要, 2-不重要, 3-普通, 4-重要, 5-非常重要, 0-缺失	观看文化类综艺节目时, 认为播出时间段的重要程度	5
18	Q15	数值型	0-4	1-从未观看, 2-一周小于1小时, 3-一周1-3小时, 4-一周3小时以上, 0-缺失	观看此类节目的频率	6
19	Q16	数值型	0-4	1-仅供娱乐, 2-传承文化底蕴, 3-寓教于乐, 4-其他,	对文化类综艺节目	7

续编码表:

				0-缺失	看法	
20	Q17.1	数值型	0-5	1-非常不满意, 2-不满意, 3-普通, 4-满意, 5-非常满意, 0-缺失	比传统类节目更具趣味性满意度	8
21	Q17.2	数值型	0-5	1-非常不满意, 2-不满意, 3-普通, 4-满意, 5-非常满意, 0-缺失	让观众更有参与感的满意度	8
22	Q17.3	数值型	0-5	1-非常不满意, 2-不满意, 3-普通, 4-满意, 5-非常满意, 0-缺失	明星入驻的满意度	8
23	Q17.4	数值型	0-5	1-非常不满意, 2-不满意, 3-普通, 4-满意, 5-非常满意, 0-缺失	内容专业性的满意度	8
24	Q17.5	数值型	0-5	1-非常不满意, 2-不满意, 3-普通, 4-满意, 5-非常满意, 0-缺失	文化传承积极作用的满意度	8
25	Q17.6	数值型	0-5	1-非常不满意, 2-不满意, 3-普通, 4-满意, 5-非常满意, 0-缺失	电视谈影响力的满意度	8
26	Q17.7	数值型	0-5	1-非常不满意, 2-不满意, 3-普通, 4-满意, 5-非常满意, 0-缺失	舞台效果的满意度	8
27	Q17.8	数值型	0-5	1-非常不满意, 2-不满意, 3-普通, 4-满意, 5-非常满意, 0-缺失	播出时间段的满意度	8
28	Q18	数值型	0-3	1-以前就会, 2-否, 3-是, 0-缺失	在日常生活中讨论与文化类综艺节目相关的话题	9
29	Q19	数值型	0-3	1-以前就会, 2-否, 3-是, 0-缺失	查阅相关资料	10

续编码表:

30	Q20	数值型	0-3	1-以前就会, 2-否, 3-是, 0-缺失	游历文化古迹(参观博物馆或名人故居等)	11
31	Q21	数值型	0-3	1-以前就会, 2-否, 3-是, 0-缺失	参与传统文化相关行业或相关活动(例如志愿者)	12
32	Q22	数值型	0-5	1-降低很多, 2-稍有降低, 3-没有影响, 4-稍有提升, 5-提升很大, 0-缺失	观看节目对民族自豪感造成的影响	13
33	Q23	数值型	0-5	1-降低很多, 2-稍有降低, 3-没有影响, 4-稍有提升, 5-提升很大, 0-缺失	观看节目对传统文化的热爱造成的影响	14
34	Q24.1	数值型	0-7	1-非常不符合, 2-不符合, 3-较不符合, 4-一般, 5-较符合, 6-符合, 7-非常符合, 0-缺失	观看前认为文化类从业者是非主流的	15
35	Q24.2	数值型	0-7	1-非常不符合, 2-不符合, 3-较不符合, 4-一般, 5-较符合, 6-符合, 7-非常符合, 0-缺失	观看前认为文化类从业者是有前途的	15
36	Q24.3	数值型	0-7	1-非常不符合, 2-不符合, 3-较不符合, 4-一般, 5-较符合, 6-符合, 7-非常符合, 0-缺失	观看前认为文化类从业者是低收入的	15
37	Q24.4	数值型	0-7	1-非常不符合, 2-不符合, 3-较不符合, 4-一般, 5-较符合, 6-符合, 7-非常符合, 0-缺失	观看前认为文化类从业者是温文尔雅	15

续编码表:				的		
38	Q24.5	数值型	0-7	1-非常不符合, 2-不符合, 3-较不符合, 4-一般, 5-较符合, 6-符合, 7-非常符合, 0-缺失	观看前认为文化类从业者是腐朽无能的	15
39	Q24.6	数值型	0-7	1-非常不符合, 2-不符合, 3-较不符合, 4-一般, 5-较符合, 6-符合, 7-非常符合, 0-缺失	观看前认为文化类从业者的工作是铁饭碗	15
40	Q24.7	数值型	0-7	1-非常不符合, 2-不符合, 3-较不符合, 4-一般, 5-较符合, 6-符合, 7-非常符合, 0-缺失	观看前认为文化类从业者的工作福利待遇高	15
41	Q24.8	数值型	0-7	1-非常不符合, 2-不符合, 3-较不符合, 4-一般, 5-较符合, 6-符合, 7-非常符合, 0-缺失	观看前认为文化类从业者是知识渊博的	15
42	Q24.9	数值型	0-7	1-非常不符合, 2-不符合, 3-较不符合, 4-一般, 5-较符合, 6-符合, 7-非常符合, 0-缺失	观看前认为文化类从业者是无私的	15
43	Q24.10	数值型	0-7	1-非常不符合, 2-不符合, 3-较不符合, 4-一般, 5-较符合, 6-符合, 7-非常符合, 0-缺失	观看前认为文化类从业者是辛苦劳累的	15
44	Q25.1	数值型	0-7	1-非常不符合, 2-不符合, 3-较不符合, 4-一般, 5-较符合, 6-符合, 7-非常符合, 0-缺失	观看后认为文化类从业者是非主流的	16
45	Q25.2	数值型	0-7	1-非常不符合, 2-不符合,	观看后认	16

续编码表:

		型		3-较不符合, 4-一般, 5-较符合, 6-符合, 7-非常符合, 0-缺失	为文化类从业者是有前途的	
46	Q25.3	数值型	0-7	1-非常不符合, 2-不符合, 3-较不符合, 4-一般, 5-较符合, 6-符合, 7-非常符合, 0-缺失	观看后认为文化类从业者是低收入的	16
47	Q26.4	数值型	0-7	1-非常不符合, 2-不符合, 3-较不符合, 4-一般, 5-较符合, 6-符合, 7-非常符合, 0-缺失	观看后认为文化类从业者是温文尔雅的	16
48	Q25.5	数值型	0-7	1-非常不符合, 2-不符合, 3-较不符合, 4-一般, 5-较符合, 6-符合, 7-非常符合, 0-缺失	观看后认为文化类从业者是腐朽无能的	16
49	Q25.6	数值型	0-7	1-非常不符合, 2-不符合, 3-较不符合, 4-一般, 5-较符合, 6-符合, 7-非常符合, 0-缺失	观看后认为文化类从业者的工作是铁饭碗	16
50	Q25.7	数值型	0-7	1-非常不符合, 2-不符合, 3-较不符合, 4-一般, 5-较符合, 6-符合, 7-非常符合, 0-缺失	观看后认为文化类从业者的工作福利好待遇高	16
51	Q25.8	数值型	0-7	1-非常不符合, 2-不符合, 3-较不符合, 4-一般, 5-较符合, 6-符合, 7-非常符合, 0-缺失	观看后认为文化类从业者是知识渊博的	16
52	Q25.9	数值型	0-7	1-非常不符合, 2-不符合, 3-较不符合, 4-一般, 5-较符合, 6-符合, 7-非常符合, 0-缺失	观看后认为文化类从业者是	16

				0-缺失	无私的	
续编码表:						
53	Q25.1 0	数值 型	0-7	1-非常不符合, 2-不符合, 3-较不符合, 4-一般, 5-较 符合, 6-符合, 7-非常符合, 0-缺失	观看后认 为文化类 从业者是 辛苦劳累 的	16
54	Q26	数值 型	0-2	1-是, 2-否, 0-缺失	持续关注 此类节目	17
55	Q27	数值 型	0-6	1-审美疲劳, 2-节目同质化 严重, 3-内容平庸, 4-节目 不专业, 5-节目情感空泛, 6-其他, 0-缺失	不会持续 关注的原 因	18
56	Q29	数值 型	0-2	1-是, 2-否, 0-缺失	将此类节 目推荐给 他人看	19
57	Q31	数值 型	0-7	1-历史文化名城、名镇、名 村, 2-方言文化, 3-中华老 字号, 4-节庆礼仪服装服 饰, 5-历史建筑, 6-民族传 统体育项目, 7-其他, 0-缺 失	节目需要 增加的 主题领域	20

附录五

访谈稿

尊敬的朋友：

您好！我们正以“文化类综艺节目”为主题展开调查活动。我们想对您展开一个访谈，我们将会对您的访谈回答进行严格保密，本次调查不用写调查者姓名，所有回答用于学术研究。您只需根据自己的实际情况进行回答。衷心感谢您的支持与配合！

（一）针对家长和老师的采访稿

1. 请问您是否有关注过《中国诗词大会》、《朗读者》、《国家宝藏》等文化类综艺节目？
2. 想问一下，您对此类节目有何看法？
3. （老师回答，家长不答）您是否会在课堂上面播放这些节目或者课后推荐学生观看这类节目？
4. （家长回答，老师不答）您是否会让您孩子去看这些节目？是您孩子主动去看，还是您建议您孩子去看？
5. 如果节目中出现了一些难懂的问题，您是否会给学生（孩子）做些讲解？
6. 观看节目之后学校（家长您）是否会组织学生（带孩子）去参观相应的景点，如博物馆、名人故居？
7. 您觉得学生（孩子）对于传统文化的热情是否有变化？有何变化？
8. 您现在或者是将来是否会把文化类综艺节目当做一种教育方式？
9. 您是否会建议您的学生（孩子）从事文化保护、文化传承相关的工作？
10. 您为什么会建议（或不建议）您的学生（孩子）从事文化保护、文化传承相关的工作？

（二）针对外国友人的采访稿

1. 请问您之前是否有观看过《汉语桥》之类的节目呢？
2. 您是否有关注过《中国诗词大会》、《朗读者》、《国家宝藏》等文化类综艺节目？

3. 想问一下，您对此类节目有何看法？您觉得这些节目有哪些异同？
4. 在您的国家，是否存在类似的节目呢？
5. 您观看节目时候，是否对中华优秀传统文化有更深入的了解呢？
6. 在观看节目的过程中，对节目内容理解程度有多少？碰到不理解的地方是否会查询相关资料？
7. 在你的同群人之间，是否会观看或者是讨论这些节目？您是否会向他人推荐呢？
8. 您是否会去参观参观相应的景点，如博物馆、名人故居？

（三）针对博物馆、名人故居、红色景点等工作人员的采访稿

1. 请问您是否有关注过《中国诗词大会》、《朗读者》、《国家宝藏》等文化类综艺节目？
2. 您对此类节目有何看法？
3. 您从事这个工作多久？平时负责的内容有哪些？
4. 自节目播出之后，参观博物馆（或名人故居或红色景点）的人流量有何变化？
5. 参观游客的年龄层次和性别有何变化？
6. 游客在参观过程中是否会询问节目相关内容？
7. 在文化类综艺节目播出之后，国家对博物馆的支持是否增加，譬如拨款的增加？如果拨款增加，那么你们打算如何支配这些款项？
8. 节目播出之后，来担当志愿者的人数或者是全职工作的人数有何变化？

附录六

程序代码

1.情感分析代码:

#一、输入数据值

```
library(readr)
```

```
train<-read_csv("C:\\Users\\acer-1\\Desktop\\市调 2018\\文本挖掘\\gjbz.csv")
```

#二、输入词典

##2.1 情感正向词，词组+打“+1”-label

```
pos<-readLines("C:\\Users\\acer-1\\Desktop\\positive.txt")##导入正向词典
```

```
weight<-rep(1,length(pos))#weight 是标签
```

```
pos<-cbind(pos,weight)
```

##2.2 情感负向词，词组+打“-1”-label

```
neg<-readLines("C:\\Users\\acer-1\\Desktop\\negative.txt")##导入负向词典
```

```
weight<-rep(-1,length(neg))
```

```
neg<-cbind(neg,weight)
```

##2.3 正、负向词组合并

```
posneg<-rbind(pos,neg)#正负词典合并（按列往下顺延），出现 NA 则输入
```

```
options(max.print=1000000)
```

```
colnames(posneg)<-c("term","weight")
```

##2.4 读入词库

```
library(rJava)
```

```
library(Rwordseg)#需在 rJava 环境下配置
```

```
library(jiebaR)
```

```
dict<-posneg[, "term"]
```

```
listDict()#查看已有词库
```

```
uninstallDict()#删除安装的词典
insertWords(dict)
insertWords(mingxin)##加入搜狗词库明星词库
```

#三、数据清洗+分词

##3.1 一、二级清洗

```
sentence<-as.vector(train$term)#文本内容转化为向量 sentence
sentence<-gsub("[[:digit:]]*", "", sentence)#清除数字[a-zA-Z]
sentence<-gsub("[a-zA-Z]", "", sentence)#清除英文字符
sentence<-gsub("\\.", "", sentence)#清除全英文的 dot 符号
train<-train[!is.na(sentence),]#清除一些空值文本（文本名）
sentence<-sentence[!is.na(sentence)]#清除对应 sentence 里面的空值（文本内容），
要先执行文本名
train<-train[!nchar(sentence)<2,]#筛选字符数小于 2 的文本
sentence<-sentence[!nchar(sentence)<2]#`nchar`函数对字符计数，英文叹号为 R 语
言里的“非”函数
```

##3.2 分词

```
x<-segmentCN(strwords=sentence)
temp<-lapply(x,length)#每一个元素的长度,即文本分出多少个词
temp<-unlist(temp)#lapply 返回的是一个 list, 所以 3 行 unlist
id<-rep(t(train[, 'ID']),temp)#将每一个对应的 id 复制相应的次数, 就可以和词汇对
应了
term<-unlist(x)#6 行将 list 解散为向量

testterm<-as.data.frame(cbind(id,term),stringsAsFactors=F)
p<-as.data.frame(posneg,stringsAsFactors=F)
#将一一对应的两个向量按列捆绑为数据框, 分词整理就基本结束了
```

##3.3 清洗停用词

```
stopword<-read.table(file=file.choose(),stringsAsFactors=FALSE)#stopword.txt  
stopword<-setdiff(stopword,posneg)  
testterm<-setdiff(testterm$term,stopword)
```

#四、关联情感词

```
library(plyr)  
testterm<-as.data.frame(testterm)  
testterm<-join(testterm,p)#将词典和分词后的词语对应
```

```
testterm<-testterm[!is.na(testterm$weight),]#去除 NA  
head(testterm)
```

#五、情感分数

##5.1 情感分数

```
data=as.numeric(testterm$weight)  
dictresult<-aggregate(data,by=list(ID=testterm$ID),sum)#按 id 名加总每条 id 的得分
```

##5.2 情感偏向

```
dictlabel<-rep(-1,length(dictresult[,1]))#id 总体得分>0,则总体倾向为正向,记为 1;  
负向记为-1
```

```
dictlabel[dictresult$x>0]<-1
```

```
dictresult<-as.data.frame(cbind(dictresult,dictlabel),stringsAsFactors=T)
```

```
posmark<-sum(dictresult$dictlabel>0);posmark
```

```
negmark<-sum(dictresult$dictlabel<0);negmark
```

```
posmark.per<-posmark/sum(posmark+negmark);posmark.per#积极情绪百分比
```

```
negmark.per<-negmark/sum(posmark+negmark);negmark.per#负面情绪百分比
```

2.词云图（分词部分略去）

```
#调用词云图包
library(wordcloud2)
#计算词频
wordsnum=table(unlist(testterm$term))
wordsnum=sort(wordsnum)#排序
#选出词频最高的 250 个
wordsnum=tail(wordsnum,55)
#画词云图
wordcloud2(wordsnum)
```

3.LDA 主题分类

##1.预处理

```
comments_segged<-testterm#读取分词结果
comments<-as.list(comments_segged)#将向量转化为列表
doc.list<-strsplit(as.character(comments),split="")#将每行文本，按照空格分开，每
行变成一个词向量，储存在列表里
term.table<-table(unlist(doc.list[2]))
term.table<-sort(term.table,decreasing=TRUE)#按照词频降序排列
term.table2<-term.table
del<-term.table<5|nchar(names(term.table))<3#把不符合要求的筛出来
term.table<-term.table[!del]#去掉不符合要求的

term.table

vocab<-names(term.table)#创建词库
get.terms<-function(x){
index<-match(x,vocab)}#获取词的 ID
```

```

index<-index[!is.na(index)]#去掉没有查到的，也就是去掉了的词
rbind(as.integer(index-1),as.integer(rep(1,length(index))))#生成上图结构}
documents<-lapply(doc.list,get.terms)

##2.LDA 建模（分词部分略去）
K<-5#主题数
G<-5000#迭代次数
alpha<-0.10
beta<-0.02
library(lda)
set.seed(357)
fit<-lda.collapsed.gibbs.sampler(documents=documents,K=K,
vocab=vocab,num.iterations=G,alpha=alpha,eta=eta,
initial=NULL,burnin=0,compute.log.likelihood=TRUE)

##3、LDAvis 可视化
theta<-t(apply(fit$document_sums+alpha,2,function(x)x/sum(x)))#文档—主题分布
矩阵
phi<-t(apply(t(fit$topics)+eta,2,function(x)x/sum(x)))#主题-词语分布矩阵
term.frequency<-as.integer(term.table)#词频
doc.length<-sapply(documents,function(x)sum(x[2,]))#每篇文章的长度，即有多少
个词

library(LDAvis)
json<-createJSON(phi=phi,theta=theta,doc.length=doc.length,vocab=vocab,term.freq
uency=term.frequency)
serVis(json,out.dir='./vis',open.browser=FALSE)
###改成 UTF-8 格式以免出现乱码
writeLines(iconv(readLines("./vis/lda.json"),from="GBK",to="UTF8"),file("./vis/lda.j

```

```
son",encoding="UTF-8"))
```

4.Ridit 分析法

```
>Ridit.pan<-c(data)#导入数据
```

```
>B<-matrix(Ridit.pan,5,5)#将数据转化为矩阵
```

```
>Ridit.test(B)#进行 Ridit 检验
```

```
>Ridit.test=function(x)
```

```
+{
```

```
+order.num=ncol(x)
```

```
+treat.num=nrow(x)
```

```
+rowsum=rowSums(x)
```

```
+colsum=colSums(x)
```

```
+total=sum(rowsum)
```

```
+N=(colsum/2)[1:order.num]
```

```
+c(0,(cumsum(colsum))[1:order.num-1])
```

```
+ri=N/total#每个顺序类的得分
```

```
+p_coni=x/outer(rowsum,rep(1,order.num),"*")
```

```
+pi.=rowsum/total
```

```
+score=p_coni%*%ri#每个处理的得分
```

```
+confi_inter=matrix(c(score-1/sqrt(3*rowsum),score+1/sqrt(3*rowsum)),byrow=F,ncol=2)#置信区间矩阵
```

```
+if(length(rle(sort(ri))$lengths)==length(ri))#不打结的情况
```

```
+{w=(12*total/(total+1))*sum(rowsum*(score-0.5)^2)}
```

```
+if(length(rle(sort(ri))$lengths)<=length(ri))#打结的情况
```

```
+{tao=rle(sort(ri))$lengths
```

```
+T=1-sum(tao^3-tao)/(order.num^3-order.num)
```

```
+w=(12*total/((total+1)*T))*sum(rowsum*(score-0.5)^2)}#统计量
```

```
+pvalue=pchisq(w,treat.num-1,lower.tail=FALSE)#p 值
```

```
+list(score=score,confi_inter=confi_inter,W=w,Pvalue=pvalue)#函数返回每个处理的得分、置信区间、统计量的值和相应 p 值
+}
>options(digits=4)#设置结果为 4 位小数
>res_B=Ridit.test(B)#运用编写的函数分析矩阵 B
>score_B=res_B$score>score_B#得到每个因素的得分
>graph_B=res_B$confi_inter
>graph_B#得到置信区间
>Pvalue_B=res_B$Pvalue>Pvalue_B#得到检验的 p 值
>plot(0,0,ylim=c(0,1),xlim=c(1,7),xlab="因素",ylab="",main="Ridit 得分置信区间图",col="gray7")#设定坐标系
>abline(h=0.5)#设定参考中横线
>for(iin1:nrow(graph_B))lines(c(i,i),graph_B[i,],lwd=2) #画出图
```

