



书舍万千，读者言欢

城市公共阅读空间现状及前景调查

——以杭州市为例

多读书读好书才队



2019 年（第六届）全国大学生统计建模大赛

书舍万千，读者言欢：城市公共阅读空间现状
及前景调查

——以杭州市为例

参 赛 单 位：浙江财经大学

参赛者姓名：胡容莉、陈昕晔、闫晋彤

目录

摘要.....	I
一、 绪论.....	1
(一) 研究背景及目的.....	1
1. 研究背景.....	1
2. 研究目的与意义.....	2
(二) 文献综述.....	2
1. 关于阅读现状的文献综述.....	2
2. 关于多维阅读空间的文献综述.....	3
3. 关于阅读空间相关政策的文献综述.....	4
4. 文献综述小结.....	5
二、 调查方案设计.....	5
(一) 调查流程.....	5
1. 调查思路.....	5
2. 技术路线图.....	6
(二) 调查方式.....	6
1. 网络预调查.....	6
2. 问卷调查.....	7
三、 质量控制.....	11
(一) 数据预处理.....	11
(二) 信度和效度检验.....	11
四、 基于 LDA 模型的独立书店评论主题挖掘.....	12
(一) 文本预处理.....	12
1. 数据说明.....	12
2. 文本分词与词频统计.....	12
(二) 杭州独立书店文本 LDA 主题模型.....	13
五、 居民阅读现状与公共阅读空间现状分析.....	16
(一) 基本信息.....	16
(二) 居民阅读现状的图表分析.....	18
1. 阅读目的分析.....	18

2. 阅读类型分析.....	18
3. 阅读媒介分析.....	19
4. 阅读数量与原因分析.....	19
(三) 影响阅读时长的多变量线性回归.....	20
(四) 影响阅读质量满意度的 Logit 定序回归.....	21
(五) 公共阅读空间现状分析.....	25
1. 居民前往阅读空间的目的分析.....	25
2. 阅读空间推广及普及.....	25
3. 读者偏好与阅读空间优势.....	26
4. 不同类型阅读空间优劣势分析.....	27
六、 公共阅读空间需求影响因素分析.....	28
(一) 阅读空间需求因素的探索性因子分析——主成分法.....	28
(二) 阅读空间需求因素验证性因子分析——PLS 路径模型.....	30
(三) 各类公共阅读空间目标群体定位——关联规则.....	32
1. 模型构建.....	33
2. 关联规则结果分析.....	34
七、 结论与建议.....	36
(一) 结论.....	36
(二) 建议.....	37
1. 宏观层面关于战略文化布局的建议.....	37
2. 中观层面关于文化空间完善的建议.....	38
3. 微观层面关于阅读氛围营造的建议.....	38
4. 各类公共阅读空间建设框架.....	39
参考文献.....	40
附录一：调查问卷.....	42
附录二：调查照片.....	48

表格和插图清单

图 1 技术路线图.....	6
图 2 抽样框设计及组织形式.....	8
图 3 前 100 个关键词词云图.....	13
图 4 独立书店评论主题气泡图.....	14
图 5 调查对象人口金字塔.....	16
图 6 文化程度与阅读目的比对图.....	18
图 7 阅读媒介评分.....	19
图 8 阅读数量少的原因对比图.....	19
图 9 前去阅读空间目的得分图.....	25
图 10 公共阅读空间推广及普及情况图.....	25
图 11 读者偏好与阅读空间优势图.....	26
图 12 不同类型阅读空间优劣势图.....	27
图 13 碎石图.....	28
图 14 PLS 路径模型图.....	31
图 15 PLS 路径模型权重系数图.....	31
图 16 公共阅读空间建设框架图.....	39
表 1 杭州市抽样阶段表.....	8
表 2 市辖区分层情况表.....	9
表 3 第一阶样本容量分配表.....	9
表 4 代码法实施表.....	10
表 5 第二阶段样本容量.....	10
表 6 最终样本容量.....	11
表 7 量表信度分析表.....	12
表 8 量表效度分析表.....	12
表 9 评论数据样例.....	12
表 10 独立书店最终评论主题提取表.....	14
表 11 样本基本信息表.....	17
表 12 被调查者阅读类型表.....	18
表 13 不同因变量多变量回归结果对比.....	20
表 14 多变量回归模型对比系数表.....	21
表 15 影响阅读质量满意度因素显著性检验表.....	22

表 16 logistic 回归系数表.....	23
表 17 logistic 回归截距表.....	23
表 18 方差贡献率.....	28
表 19 旋转成份矩阵.....	29
表 20 因子载荷矩阵.....	30
表 21 测量模型检验表.....	31
表 22 结构模型检验表.....	32
表 23 PLS 路径模型效应表.....	32
表 24 前项数据预处理表.....	33
表 25 惠民便捷因子关联分析表.....	34
表 26 拓展阅读因子关联分析表.....	35
表 27 附加价值因子关联分析表.....	36

摘要

在信息化时代背景下，人们的浅阅读、碎片式阅读现象普遍存在，因此创建完善的城市阅读空间体系是引导大众阅读、构建我国文化服务体系中极为重要的一环。但在国家政策引导和各地政府支持下，城市阅读空间的发展过程仍然存在着诸如高成本自助借还机遇冷、实体书店纷纷倒闭等问题。为此，本次调查旨在通过问卷调查与数据挖掘，探究社会阅读现状及公共阅读空间发展现状，以读者视角为出发点，提出针对性建议，使阅读空间的经营方向切合读者的阅读需求。

项目组首先采用多阶段分层抽样对杭州市居民进行抽样；通过 LDA 主题建模对豆瓣独立书店评价日记进行主题归类，了解人们对于独立书店的主要评价角度，探析当今社会阅读现状并对问卷进行适当调整。在实地发放问卷后，运用图表分析居民阅读现状、公共阅读空间发展现状；建立多变量线性回归探究影响居民阅读时长的因素；建立 Logit 定序回归模型，探究影响阅读满意度的因素，根据模型进行读者满意度预测；运用探索性因子分析提取影响公共阅读空间阅读体验的因子；建立 PLS 路径模型，探究各因子的直接与间接影响；用 Apriori 关联规则挖掘算法为不同类型阅读空间定位目标客户群体；综合以上分析提出针对性建议。

研究结论如下：文化素养与阅读需求、阅读时长成正向关系，年轻的学生群体是前往阅读空间的主力军；电子阅读媒介普及广泛，公共阅读体系无法延伸到公民日常生活中，导致碎片化阅读、浅阅读现象普遍存在；公共阅读空间可以通过提供阅读环境和营造阅读氛围促进阅读，但不应只局限于提供阅读资源，与此同时公共阅读空间还存在缺乏宣传的问题；“互联网+”推动公共阅读空间进一步发展；不同阅读空间存在不同局限性，目标群体不同；环境因素对公共阅读空间有显著影响。最后，根据研究结论，从宏观、中观、微观三个层面对公共阅读空间未来发展提出针对性建议，构建各类公共阅读空间建设框架

关键词：公共阅读空间；LDA 主题模型；关联分析；PLS 路径模型

一、 绪论

(一) 研究背景及目的

1. 研究背景

在数字化、信息化的时代背景下，没时间读书、阅读碎片化、快餐化、娱乐化、功利化等已成为一种普遍现象。我国第十六次《全民阅读调查报告》表明，2018年我国国民人均纸质图书阅读量为4.67本，与2017年的4.66本相比并没有明显提高，并且与日本国民人均纸质阅读数量约为11本，美国人均7本等相比较，当今国人的阅读状况并不乐观。

由政府出资的，我国最早开展地铁图书馆的实践可以追溯到2008年，随着城市轨道交通建设加快，北京、深圳、东莞、长沙、宁波等城市都开始了地铁图书馆的建设。上海地铁图书馆在最开始的一周内，平均每天流失图书500多本，图书流失率高达50%。武汉地铁自助图书馆运行10个月的数据显示，累计借书8万余册，其中1万余册书没有及时归还，失书现象严重。并且一台自助借还机引进成本高达40万，一条地铁线至少有25个以上的站点，建设成本之高让人却步。与此同时自助借还机并未得到广大市民的普遍接受，自助借还机遇冷的新闻频频出现在各大媒体。

我国城市民营独立书店经营转型的探索过程中也在不断发展公共阅读空间。但王子舟（2017）指出，自2011年以来受书店租金上涨、高税收与人员成本增长、电子书流行、图书网购普及等因素的冲击，许多实体独立书店因保持不住微薄的利润而纷纷倒闭。例如2010年北京“第三极”书局倒闭，2011年北京风入松书店、北京光合作用书店、成都大印象书局，2012年上海万象书店、成都时间简史书坊都先后倒闭。据统计，从2007-2011年，中国倒闭的民营书店有一万多家。

面对我国国民阅读数量少、缺乏良好阅读习惯的现状，政府和企业都采取过一些措施来创造更好的阅读空间以及更为广泛的借阅方式，但现实情况往往无法达到预期。本文将在接下来的调查及其结果的分析中进一步探究这其中存在的问题。

2. 研究目的与意义

本文将探究现当代人们的阅读习惯和特点，通过对现存的各种公共阅读空间的经营状况和空间特点的研究，了解其作为阅读空间在激发阅读兴趣和提高深入阅读方面的作用。通过对读者的问卷调查以及文献资料的查询，探究不同类别的公共阅读空间现存的问题，以读者的视角为出发点，为提高各类公共阅读空间的有效性提出建议和改进方案，以达到阅读空间的经营方向切合读者的阅读需求的目的，为读者提供更加舒适的环境和更加丰富的图书资源，同时也能吸引更多的希望可以多阅读却因种种原因无法做到的潜在读者，从而达到促进全民阅读的结果。

（二） 文献综述

1. 关于阅读现状的文献综述

第十六次全国《国民阅读调查报告》显示，2018年我国成年国民人均纸质图书阅读量为4.67本，相比2017年我国成年国民人均纸质图书阅读量4.66本，2016年4.65本，2015年4.58本，2014年4.56本，近些年我国国民的阅读数量并没有明显提升。日本国民人均纸质阅读数量约为11本，韩国人均9本，美国人均7本，法国人均8本，中国国民阅读量明显较少。

而未成年人在这方面有更好表现：0-17周岁未成年人的人均图书阅读量为8.81本。未成年人阅读一直是全民阅读推广的重中之重，但本次国民阅读调查显示，2018年我国0至17周岁未成年人图书阅读率为80.4%，低于2017年的84.8%。对于未成年人阅读率下降的问题，中国新闻出版研究院国民阅读研究与促进中心主任徐升国认为：近年来，数字阅读、有声阅读，在未成年人群体，甚至在0至8周岁儿童阶段，已经越来越普及，这也导致了纸质阅读的减少。

第十六次全国《国民阅读调查报告》还提到：我国成年国民人均每天手机接触时长为84.87分钟，比2017年的80.43分钟增加了4.44分钟；人均每天互联网接触时长为65.12分钟，比2017年的60.70分钟增加了4.42分钟。我国成年国民网上活动行为中，以阅读新闻、社交和观看视频为主，娱乐化与碎片化特征明显，深度图书阅读行为的占比偏低。38.4%的成年国民更倾向于“拿一本纸质图书阅读”，比2017年的45.1%下降了6.7个百分点；40.2%的国民倾向于“手机阅读”，比2017年的35.1%上升了5.1个百分点。数字化阅读正以惊人的速度不断创新，也是未来大的方向与趋势。数字化阅读丰富了人们的阅读内容、场景、

方式，也提升了国民的阅读率、阅读量与阅读时长。但它也有弊端：打乱、切断了整块时间，挤占人们对深度内容进行阅读的时间；大量碎片化的阅读容易导致信息过载、爆炸，使人难以形成系统化的知识结构和判断力。

2. 关于多维阅读空间的文献综述

王炎龙等（2018）认为，公共阅读空间是由政府、企业、社会组织或个人创办，以文献资料为主要载体向公众提供知识服务的开放空间和场所，具义上包括图书馆、书城、独立书店、社区书屋等能够进行阅读交往的空间实体。

根据阅读空间的主办方、地域分布、开立原因、目标人群等因素，以及相关的文献资料，这里将阅读空间主要分为以下四个类别。

（1）基于公共图书馆的文化共享空间

叶帆（2018）指出，公共图书馆是一种主要由政府资助和支持，为所有社会成员平等免费提供信息知识服务的大型阅读空间。作为最主要的公共阅读空间之一，它具有空间大，馆藏丰富、全面的特点。

但公共图书馆常常由于公共阅读资源过于集聚而无法实现阅读日常化。并且我国公共图书馆功能布局单调，缺乏对公众的吸引力。

（2）基于社区书屋的文化便民空间

在大型公共图书馆无法满足大多数人的日常阅读需求的情况下，催生了以社区为单位的阅读空间。叶帆（2018）认为，社区书屋作为公共图书馆微型化、迷你化的版本，是当地文化部门在社区设立的一种面向公众开放的公共图书室。社区书屋的配置与居民的生活社区想联系，可以提供更为便利、快捷的文化服务，满足市民娱乐、学习、交往等多方面的需求。扩大了图书馆的服务时间和半径，是对公共图书馆功能的有益补充。

但社区书屋在长期的建设过程中，同样暴露出经费不足、藏书质量不高、服务标准欠缺等低效问题。

（3）基于独立书店的文化地标空间

以独立书店为代表的建立在繁华闹市区的城市新型阅读空间逐渐成为公共文化服务的补充力量。刘雅薇（2015）指出，独立书店既是城市文化生活的聚集地，也是一个城市最重要的风景线与文化地标，在功能布局上，独立书店集合了餐饮、文创、展览等业态，重在向为消费者打造独特体验空间的复合型空间转变，体现了多元化的布局思维。

但目前也有一些书店以销售饮品、文创等为主要项目，图书则沦为书店环境装饰的符号和制造情调的附属品。

(4) 基于智慧书店的文化创新空间

随着互联网技术、大数据、云计算、人工智能等技术的深入运用，以共享书店、无人售货书店、自动借书机等为代表的创新型阅读空间开始出现。

但成本过高，并且大众的接受程度较低，往往会出现被闲置的情况。想要进一步向社会普及还需要技术和书店更好的契合度。

3. 关于阅读空间相关政策的文献综述

2014 年“倡导全民阅读”首次写入政府工作报告。

2015 年国务院相继转发颁布了《关于做好政府向社会力量购买公共文化服务工作的意见》、《关于在公共服务领域推广政府和社会资本合作模式指导意见的通知》两份重要文件，各地政府共建城市公共阅读空间成为推广全民阅读的创新举措。

2016 年 6 月，中宣部、国家新闻出版广电总局等 11 部委还联合下发了《关于支持实体书店发展的指导意见》，鼓励实体书店创新经营模式，积极参与公共文化服务建设。12 月 25 日，全国人民代表大会公布了《中华人民共和国公共文化服务保障法》，27 日，国家新闻出版广电总局发布了《全民阅读“十三五”时期发展规划》。

《关于开展 2018 年全民阅读工作的通知》明确提出，要积极推进城市书吧、社区书屋、职工书屋等新型阅读空间建设。

这些文件释放出了国家积极推进公共文化服务事业与全民阅读活动纵深发展的意愿。在国家意志的推动和社会各行业的参与下，可预见我国公共阅读空间的发展将很快迎来个繁荣的局面。

4. 文献综述小结

1. 虽然政府从 2014 年开始重视全民阅读，但是我国人均纸质阅读数量并没有发生显著改变。

2. 从公共阅读空间研究现状来看，学者对公共阅读空间的研究多停留在理论阐述，缺少一定的数据支持；研究多位于宏观层面，少有探究读者的切实需求；研究内容上多从公共阅读空间的某一特定类型出发，缺少多维阅读空间横向比较。

二、 调查方案设计

（一） 调查流程

1. 调查思路

通过查阅相关资料初步设计问卷，使用文本挖掘对我国阅读空间发展现状进行初步分析并对问卷内容作出修改；分析问卷数据，分别对居民阅读现状及公共阅读空间发展现状进行分析，在此基础上通过多变量线性回归模型、有序 logistic 回归模型、探索性与验证性因子分析、关联规则对问卷数据进一步具体分析。最后，综合数据分析结果及相关文献资料，提出相关建议。

2. 技术路线图

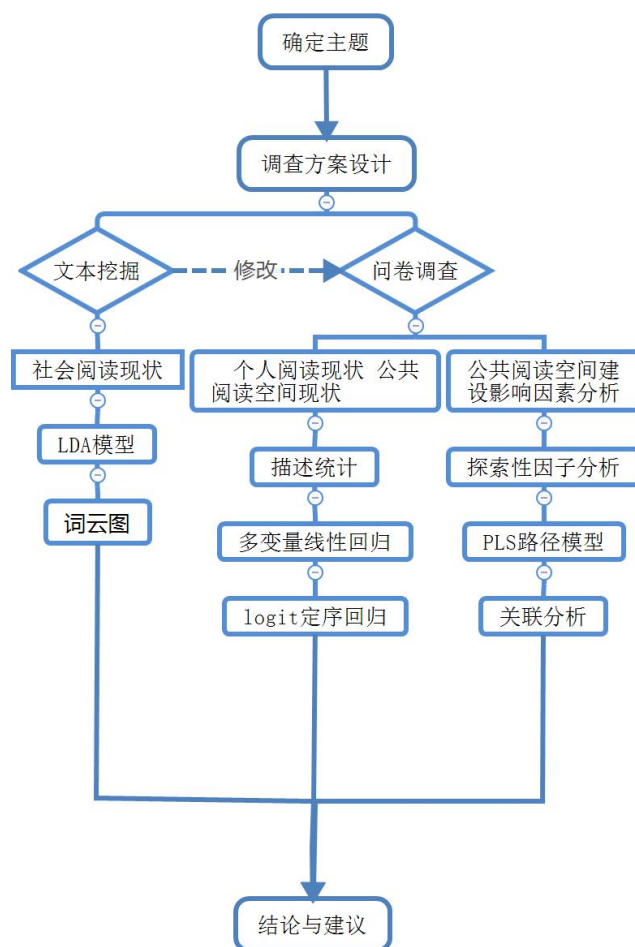


图 1 技术路线图

(二) 调查方式

本次调查主要采取网络预调查和问卷调查相结合的方式,以获得更为完善的公共阅读空间与读者现状信息作为后续分析的数据基础。

1. 网络预调查

预调查共发放并回收问卷 120 份,有效问卷 114 份,有效率为 95%。通过网络预调查,希望从中得知本次调查问卷的大致有效率,对阅读空间发展现状形成初步认知,并根据其所得数据对问卷题目与选项的设置进行合理调整。

2. 问卷调查

本次实地问卷调查首先通过查阅 2018 年杭州市统计年鉴，获得截止 2017 年末杭州市十大区各区总人口数，为使样本更具科学性与代表性，结合多种抽样方式对杭州市民进行抽样，分别运用分层比例抽样、PPS 抽样、等额抽样及偶遇抽样确定具体问卷发放地点和样本容量。

①确定样本容量

取置信度为 95%， $\mu_{\frac{\alpha}{2}}^2 = 1.96$ ，绝对误差限度 $d = 0.05$ ，总体方差 $S^2 = p(1 - p)$ ，由于成数 p 无法由先前资料得到，所以取 $p = 0.5$ 。

据统计，截止 2017 年末杭州市十个市辖区总人口数为 6152286 人。

依据总人口数 $N=6152286$ ，计算样本容量：

$$\begin{aligned}n_0 &= \frac{N\mu_{\frac{\alpha}{2}}^2 S^2}{Nd^2 + \mu_{\frac{\alpha}{2}}^2 S^2} = \frac{N\mu_{\frac{\alpha}{2}}^2 p \times (1 - p)}{Nd^2 + \mu_{\frac{\alpha}{2}}^2 p \times (1 - p)} \\&= \frac{6152286 \times 1.96^2 \times 0.5 \times 0.5}{6152286 \times 0.05^2 + 1.96^2 \times 0.5 \times 0.5} \\&\approx 384.1 \rightarrow 384\end{aligned}$$

根据设计效应为 1.35 对样本量进行调整：

$$n_1 = deff \times n_0 = 384 \times 1.35 = 518.4$$

考虑到抽样过程中会出现无法回答（包括无效问卷）等情况，假定问卷回收率达到 95%，粗略估计样本容量为：

$$n = \frac{n_1}{r} = \frac{518.4}{0.95} \approx 545.7 \rightarrow 546$$

综合实际情况，初步确定分发问卷 550 份（根据实际发放情况修改）以满足调查需要。

②抽样框设计及抽样组织形式

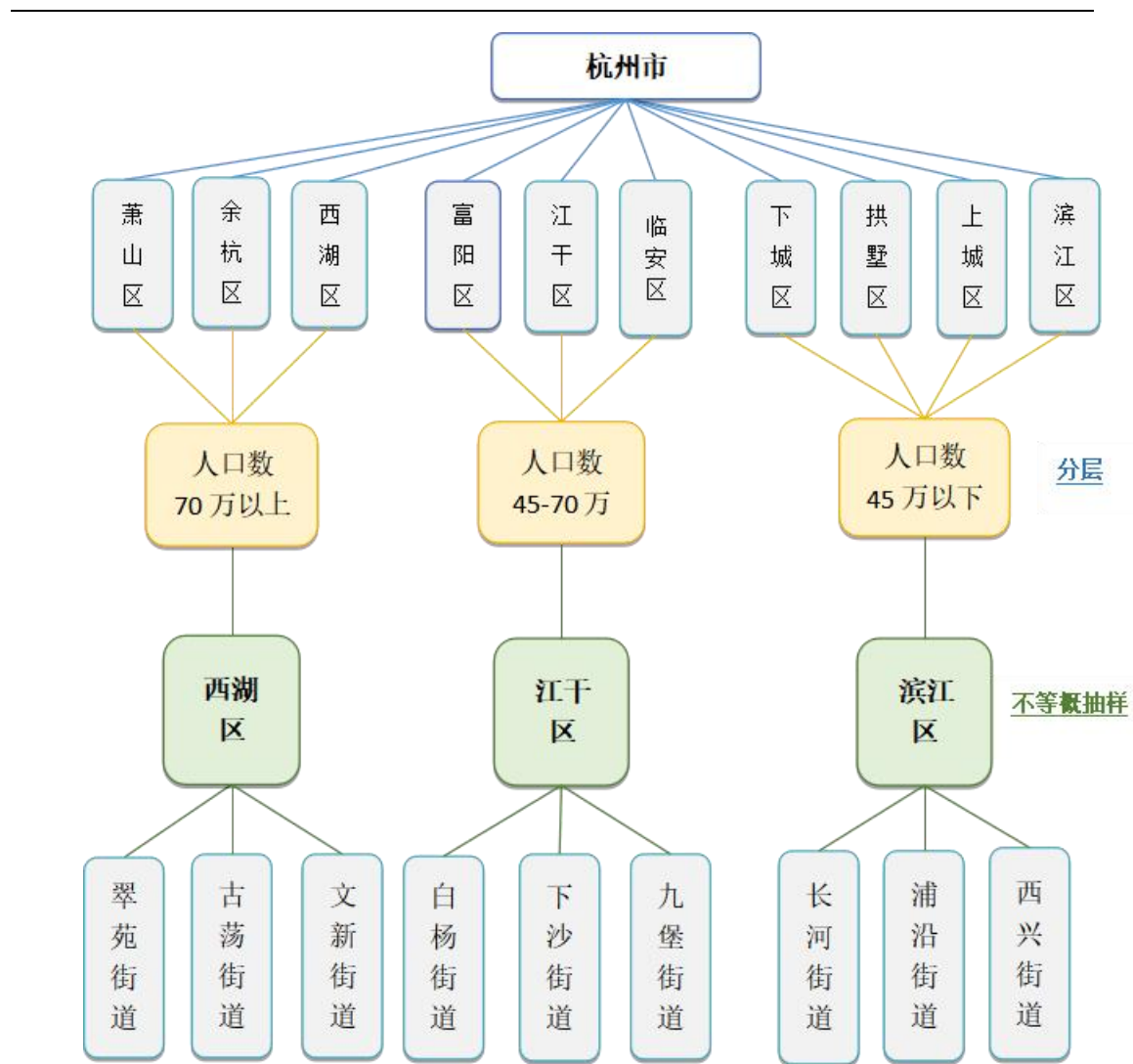


图 2 抽样框设计及组织形式

因本次在杭州市所做的公共阅读空间调查对象为本市市民，为使样本更具代表性，选择杭州市十个市辖区，并细分至街道，最终对街道居民进行调查。采用多阶段随机抽样法，按照各区人口数分阶段进行抽样，如表所示：

表 1 杭州市抽样阶段表

阶段	抽样单元	抽样方法
第一阶段	杭州市 10 个市辖区	分层比例抽样
第二阶段	第一阶段各层	PPS 抽样
第三阶段	第二阶段各区街道	等额分配
第四阶段	街道居民	偶遇法

第一阶抽样单元：

将杭州市 10 个市辖区按 2017 年末各区总人口数分为三层：

表 2 市辖区分层情况表

层数	总人口数区间	市辖区
第一层	总人口 70 万以上	萧山区、余杭区、西湖区
第二层	总人口 45-70 万	富阳区、江干区、临安区
第三层	总人口 45 万以下	下城区、拱墅区、上城区、滨江区

各市辖区的具体总人口数量与样本容量分配如下：

表 3 第一阶样本容量分配表

市辖区	总人口数	人口比例	样本容量
萧山区	1295770	21.06%	116
余杭区	1040496	16.91%	93
西湖区	712763	11.59%	64
富阳区	677210	11.01%	61
江干区	562880	9.15%	50
临安区	535154	8.70%	48
下城区	403170	6.55%	36
拱墅区	361609	5.88%	32
上城区	320431	5.21%	29
滨江区	242803	3.95%	22

第二阶抽样单元：

第二阶段运用计算机按市各辖区人口数所占总人口数的比例，在各层中随机抽取若干市辖区样本。

考虑到不同样本市辖区居民存在一定差异，为提高估计精度，减少抽样误差，在各层中用在各市辖区人口数量规模度量，用代码法实施放回不等概抽样。以第一层为例，具体操作如下：赋予每个单元与相等代码数，并将其与的百分比进行累加，如表：

表 4 代码法实施表

I	M	累计百分比
1 (萧山区)	1295770	0.4250
2 (余杭区)	1040496	0.7663
3 (西湖区)	712763	1.0000
评论总数	3049029	

实际操作中，使用 R 软件随机产生 5 个四位的随机数，设为 m，则 m 所对应的单元被抽中。

```
>floor(runif(1, 1000, 9999))
[1] 8376
>floor(runif(1, 1000, 9999))
[2]4153
>floor(runif(1, 1000, 9999))
[3] 9439
```

实际操作中：所选区为西湖区、江干区、滨江区。

使用比例分配确定样本容量如下：

表 5 第二阶段样本容量

层数	总人口数区间	市辖区	人口比例	样本容量
第一层	总人口 70 万以上	西湖区	46.94%	258
第二层	总人口 45-70 万	江干区	37.07%	204
第三层	总人口 45 万以下	滨江区	15.99%	88

第三阶抽样单元：

通过查阅资料得到第二阶抽样结果的三个市辖区所包含的街道社区，并对各层所包含街道都进行简单随机抽样：

1. 将各样本区所有的街道社区排序编号；
2. 采用随机数法，利用计算机从各样本区随机抽取 3 个样本街道社区。

通过上述步骤确定每个市辖区所抽取的街道社区后，对抽取到的街道社区进行等额分配，最终确定每个街道的样本量如下表所示：

表 6 最终样本容量

层数	市辖区	街道社区	样本容量
第一层	西湖区	翠苑街道	86
		古荡街道	86
		文新街道	86
第二层	江干区	白杨街道	68
		下沙街道	68
		九堡街道	68
第三层	滨江区	长河街道	30
		浦沿街道	30
		西兴街道	30

第四阶抽样单元:

在得到第三阶段抽样的结果后,将实地前往上述街道,采用偶遇法对所抽街道附近居民进行问卷访问,并佐以网络问卷。

三、 质量控制

(一) 数据预处理

对收集的问卷进行数据录入前,先对问卷信息及相关数据进行审核,剔除有明显矛盾和错误、或者问卷各题所选项按照某种规律进行等无效问卷。

根据抽样设计的要求,本次调查共发放问卷 580 份,回收有效问卷 543 份,问卷有效率约为 93.62%;调查对象男女比例约为 1.00:1.02,所得数据可以进行描述性分析并建立相关模型。

(二) 信度和效度检验

由于调查问卷项目设置详细,涵盖被调查者的各种信息,出于对后期数据分析及建立模型的科学性和准确性考虑,选择着重对量表项目进行信度和效度检验。选择常用的 Cronbach's Alpha 系数来检验可信度,检验结果显示量表系数值大于 0.9。就量表总体而言,具有很高的内在一致性,可靠性较强。

表 7 量表信度分析表

Cronbach's Alpha	项数
.954	20

在因子分析中，数据 KMO 值大于 0.9，Bartlett 球形检验为 9381.562，显著水平平均达到 0.000。结果表明，量表总体效度较高，适合做因子分析。同时，对问卷的各项进行了分析讨论，因而也保证了较好的内容效度。

表 8 量表效度分析表

取样足够度的 Kaiser-Meyer-Olkin 度量。		0.946
Bartlett 的球形度检验	近似卡方	9381.562
	df	190
	Sig.	0.000

四、基于 LDA 模型的独立书店评论主题挖掘

(一) 文本预处理

1. 数据说明

本文应用至 LDA 模型的数据来自于国内知名社区型评论网站豆瓣网上的评价数据。选用的数据为全国范围内独立书店的评价日记，均为用户前往过独立书店后对于该书店的评价信息。为方便分析独立书店整体状况，删除部分冗余数据，例如用户 ID、评价时间等，仅留下评论的主体文本内容。

利用集搜客软件爬取数据，具体爬取时间为 2019 年 5 月 20 日，经过初步整理，剔除重复以及无效评论后，最终可用评价文本字数为 450239 字。数据样例如下表所示：

表 9 评论数据样例

评论 1。	据说，有些节日是为严重缺乏濒危到需要提醒或提倡的人或事而设立的，比如.....(略)。
评论 2。	香港有一间独立书店叫“MOSES”，意指青苔，位置也像青苔.....(略)。
...	...

2. 文本分词与词频统计

在对豆瓣独立书店评论的实际处理应用中，选用的分词词库词汇容量约为 29.8 万；为使最终文本数据丢失较少的语义，在此词库基础上补充了搜狗词库

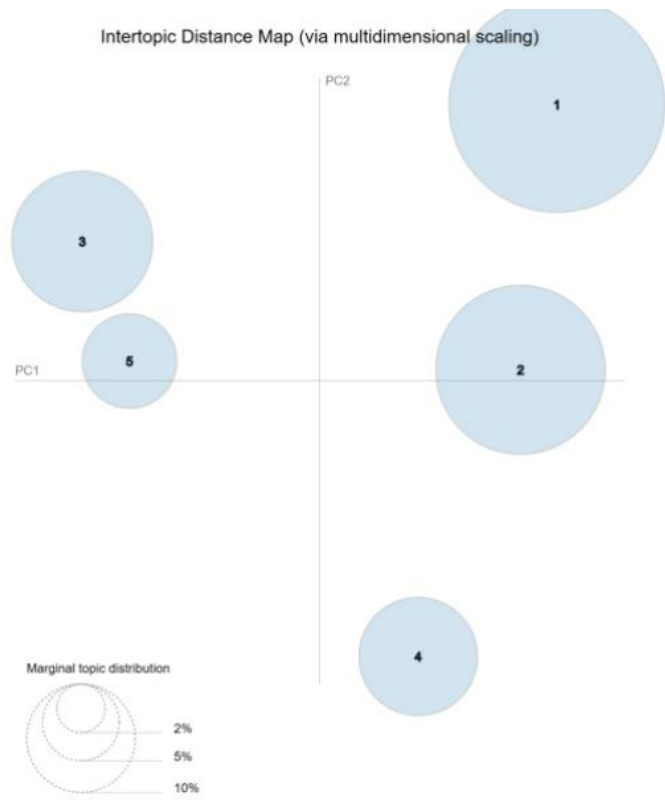


图 4 独立书店评论主题气泡图

在该评论文本数据中，主题 1 出现频率最高，主题 5 出现频率最低。且由上图可以看出，各个主题气泡并没有出现重叠部分，分布较为分散。

通过与先前同样的方法选取出每个主题下生成概率排名前 10 的词语列表，如下表所示：

表 10 独立书店最终评论主题提取表

Topic1	Topic2	Topic3	Topic4	Topic5
书店	阅读	文化	图书馆	收藏
自己	咖啡	设计	读者	丰富
可以	书籍	中国	预约	密码
看到	体验	工作	艺术	城市
地方	世界	空间	中心	新书
觉得	创意	文字	时间	市场
书架	读书	方式	志愿者	故事
喜欢	生活	建筑	公益	现场
看书	历史	未来	公众	小镇
发现	文化	分享	开放	书友

由上表看出，各个主题有明显实际意义的区分，下面对各个主题进行分类总结。

主题 1 是读者对于独立书店的主观感受,同时结合主题 1 出现的频率是最高的,与平时浏览评论网站时所观察到的现象一致。

主题 2 既有人们的深层感悟,也是关于独立书店衍生服务的评价。部分独立书店通过“网红”性质,典型如“咖啡”一类的餐饮增值服务,还有不断推出阅读相关主题活动来吸引顾客注意力,这类内容都能够显著提高一个独立书店的话题度,成为现代人们评价阅读空间的重要角度。

主题 3 则是关于独立书店整体装修设计方面的评价。人们对于一家书店的第一印象绝大部分是由其装修设计所决定的,因此日前商业化的独立书店对于设计风格均有一定的诉求,例如目前“钟书阁”的镜面设计,杭州“晓书馆”的高耸书架,等等。别具一格的装修设计能够更好的展现出书店的独特性,提升书店的口碑。

主题 5 是关于书店的藏书方面评价。一家书店的藏书是否丰富代表其实用性,如果进入这家书店的读者是抱有阅读的目的而来,那么藏书将会是其一个重要的关注点;但主题 5 气泡在图中的面积最小,这也在一定程度上反映出了部分人们进入书店并不是出于阅读目的。

主题 4 则是关于书店的公益性以及开放形式。从主题内容中可以看到许多独立书店会涉及“预约”,例如杭州的人气书店晓书馆,若要前往需要提前预约名额;在实地调查中发现一些阅读空间是需要付费进入的,因此独立书店的“公开”程度以及其“公益性”也是人们的一个讨论热点。

五、居民阅读现状与公共阅读空间现状分析

(一) 基本信息

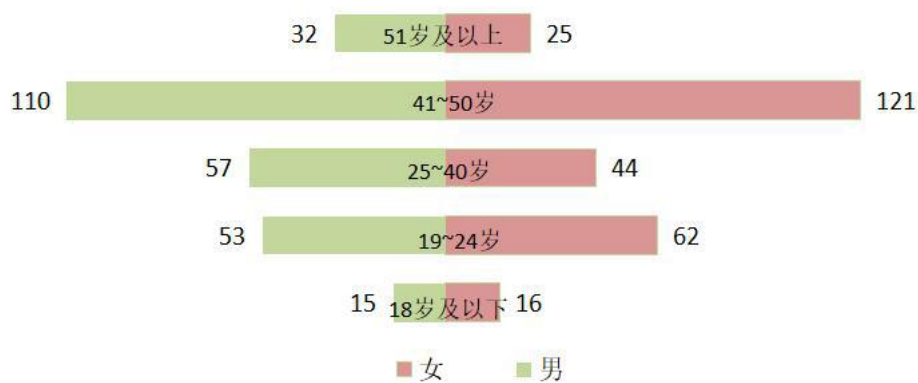


图 5 调查对象人口金字塔

本次调查回收有效问卷 543 份。调查对象男女比例约为 1: 1.02，男女比例较为平衡。其中 41~50 岁的被调查者占比较大。

表 11 样本基本信息表

	变量	人数	百分比 (%)
文化程度	初中及以下	20	3.7
	高中或中专	50	9.2
	大专	106	19.5
	大学本科	275	50.6
	研究生及以上	92	16.9
职业	学生	150	27.6
	企业管理者	29	5.3
	企业一般工作人员	56	10.3
	机关事业单位领导干部	20	3.7
	机关事业单位普通工作人员	86	15.8
	教师	71	13.1
	医生	33	6.1
	服务人员	9	1.7
	军警	17	3.1
	个体户	22	4.1
	自由职业者	1	.2
	农民	13	2.4
	农民工	13	2.4
	离退休人员	11	2.0
	下岗失业人员	1	.2
	其他	1	.2
	月收入	无收入	124
3000 元及以下		85	15.7
3001~5000 元		121	22.3
5001~8000 元		125	23.0
8001~10000 元		52	9.6
10000 元以上		36	6.6

由表 5.1 可知, 样本数据中, 文化程度占比最大的是大学本科, 占比 50.6%, 其次是大专, 占比 19.5%。样本在各职业均有所分布, 其中学生占比最大, 达到 27.6%, 其次是机关事业单位普通工作人员, 占比 15.8%。高学历的学生群体通常是前往公共阅读空间的主要群体, 因此样本基本符合调查需要。从月收入来看, 占比最大的是 5001~8000 元, 其次是 3001~5000 元, 其占比均高于 20%。

综上所述, 样本的基本情况符合实际情况, 且具有较好的代表性和研究意义, 值得进一步研究。

（二） 居民阅读现状的图表分析

1. 阅读目的分析

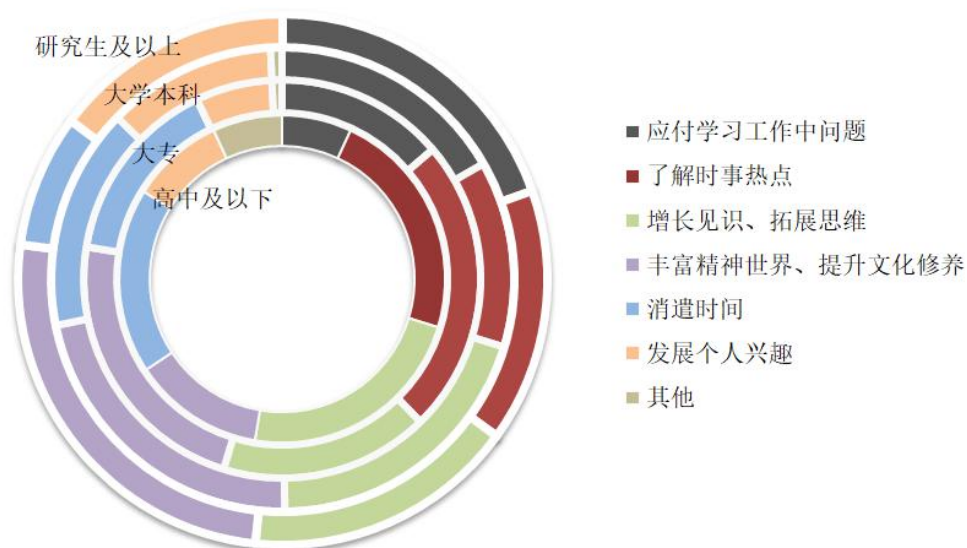


图 6 文化程度与阅读目的的比对图

由图可知，高中及以下学历的被调查者以应付学习工作中问题和丰富精神世界、提升文化修养为目的的阅读占比最小；大学本科及以上学历的被调查者阅读目的较为均衡。通过对比不同文化程度被调查者持有不同阅读目的发现：文化传播需要因人而异，阅读空间的发展应考虑不同人的文化需求。

2. 阅读类型分析

表 12 被调查者阅读类型表

哲学类	社会科学类	自然科学类	文学类	教育类	生活类	新闻热点	杂志期刊	漫画图集	其他
4.91%	11.51%	12.81%	16.03%	7.17%	13.77%	17.10%	10.38%	3.89%	2.43%

由上表可知，被调查者对新闻热点、文学类读物更感兴趣，喜好的人数占比均在 16% 以上。而对于漫画图集、哲学类读物感兴趣程度不强。

3. 阅读媒介分析

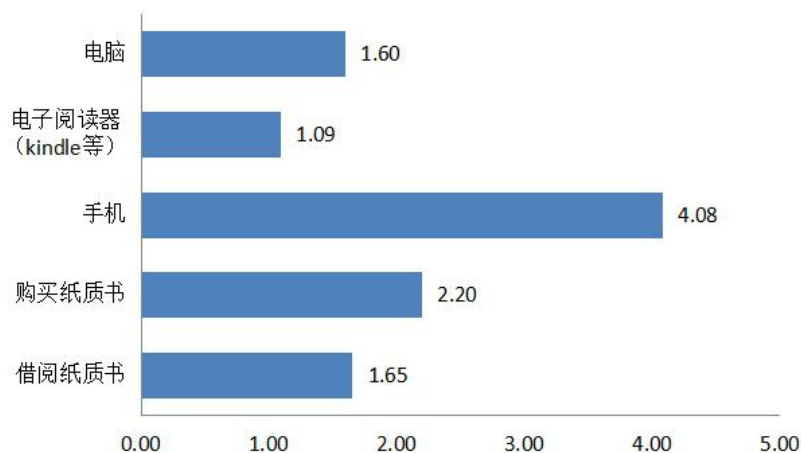


图 7 阅读媒介评分

由图可知，手机得分最高且明显高于其他媒介，说明被调查者多用手机进行阅读，但传统阅读并没有完全被线上阅读所替代。

4. 阅读数量与原因分析

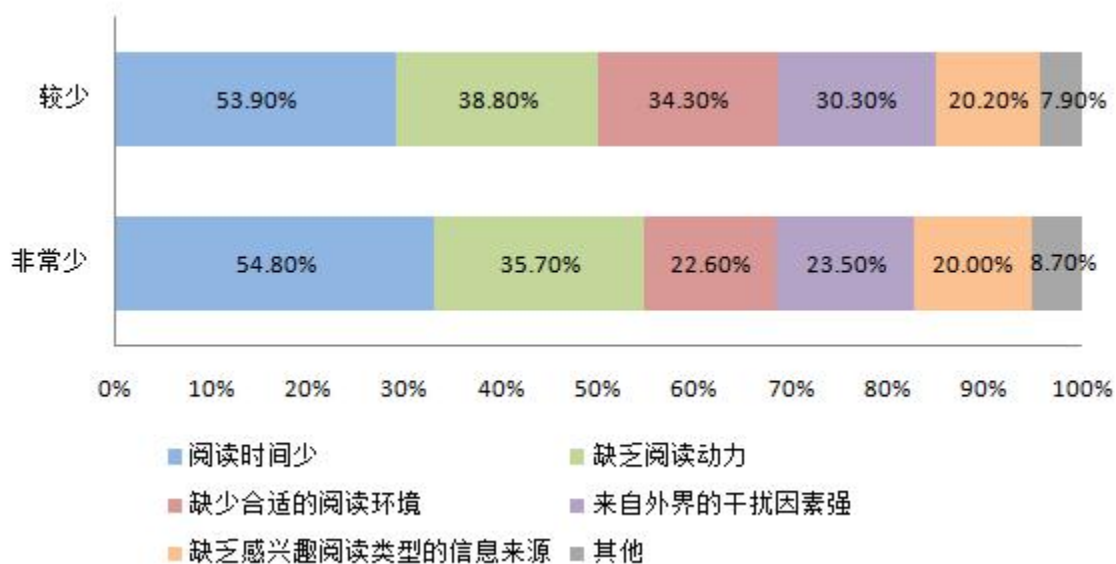


图 8 阅读数量少的原因对比图

由图可知，阅读数量非常少亦或是较少的被调查者多为阅读时间少。缺乏阅读动力也是两类人群阅读数量少的较为重要的原因。

(三) 影响阅读时长的多变量线性回归

为探究影响阅读时长的主要因素，分别以被调查者的阅读总时长（线上+线下）以及其实体书阅读时长作为因变量，筛选出以下 14 个自变量，构建多变量线性回归。得到两个模型的结果对比如表：

表 13 不同因变量多变量回归结果对比

因变量	P 值（阅读总时长）	P 值（实体书阅读时长）
性别	0.000	0.082
年龄	0.000	0.000
文化程度	0.000	0.000
月收入	0.000	0.000
应付学习或工作	0.000	0.000
了解时事热点	0.000	0.755
增长见识	0.000	0.571
丰富精神世界	0.000	0.091
消遣时间	0.063	0.047
发展个人兴趣	0.010	0.008
最常用阅读媒介	0.000	0.115
阅读数量自评	0.000	0.000
阅读质量满意度	0.000	0.091
阅读场所倾向	0.094	0.034

通过上表可知，在以阅读总时长为因变量的模型中，14 个变量的 p 值均小于 0.1，均可认为对于阅读总时长有显著影响；在以实体书阅读时长为因变量的模型中可以看到，了解时事热点、增长见识以及最常用阅读媒介三者对于被调查者实体书阅读时长没有显著影响。据此可以推断，被调查者往往不会使用实体书的方式来了解时事热点，人们通常会选择使用线上阅读的方式来了解时事热点。

接下来对两个多变量回归模型进行对比探究，其中实体书阅读市场将会去掉上述几个不产生显著影响的变量，具体回归对比结果表如下所示：

表 14 多变量回归模型对比系数表

自变量 _o	总时长 _o	实体书 _o	自变量 _o	总时长 _o	实体书 _o
截距 _o	17.79 _o	16.17 _o	(消遣时间)是 _o	1.59 _o	-0.66 _o
(性别)女 _o	-2.26 _o	0.85 _o	(发展个人兴趣)是 _o	1.35 _o	1.28 _o
(年龄)19~24 岁 _o	-7.84 _o	-11.74 _o	电子阅读器 _o	1.25 _o	- _o
(年龄)25~40 岁 _o	-7.37 _o	-10.77 _o	购买纸质书 _o	-2.42 _o	- _o
(年龄)41~50 岁 _o	-10.25 _o	-12.06 _o	借阅纸质书 _o	-3.86 _o	- _o
(年龄)51~60 岁 _o	-13.97 _o	-9.98 _o	手机 _o	2.52 _o	- _o
(文化程度)大学本科 _o	0.80 _o	-0.10 _o	(阅读数量自评)非常多 _o	42.62 _o	9.69 _o
(文化程度)大专 _o	0.34 _o	0.29 _o	非常少 _o	-11.21 _o	-6.49 _o
(文化程度)高中或中专 _o	-2.28 _o	0.66 _o	较多 _o	5.40 _o	6.67 _o
(文化程度)研究生及以上 _o	6.89 _o	1.31 _o	较少 _o	-6.58 _o	-5.18 _o
(月收入)3000 元及以下 _o	1.10 _o	1.12 _o	一般 _o	-2.76 _o	-0.72 _o
(月收入)3001~5000 元 _o	2.20 _o	1.66 _o	(阅读质量满意度)不太满意 _o	1.46 _o	0.34 _o
(月收入)5001~8000 元 _o	0.71 _o	-0.20 _o	非常不满意 _o	2.49 _o	2.05 _o
(月收入)8001~10000 元 _o	5.32 _o	5.83 _o	非常满意 _o	6.81 _o	8.09 _o
(月收入)无收入 _o	-0.29 _o	0.12 _o	一般 _o	6.07 _o	1.73 _o
(应付学习或工作)是 _o	6.81 _o	2.84 _o	(阅读场所倾向) 附加增值服务的阅读室 _o	-4.10 _o	-2.18 _o
(了解时事热点)是 _o	4.64 _o	- _o	其他 _o	-4.05 _o	-1.70 _o
(增长见识)是 _o	-6.22 _o	- _o	私人场所 _o	-2.56 _o	-1.43 _o
(丰富精神世界)是 _o	5.06 _o	0.37 _o	无其他服务的阅读室 _o	-2.74 _o	3.41 _o

由表可知模型以性别男、年龄为 18 岁及以下的文化程度为初中及以下的被调查者为基准水平建立模型，得到以下结论：

1. 相较于男性，女性的总阅读时长较短；而在实体书阅读方面，女性的阅读时长较长。随着年龄的增长，读者的总阅读时长大体上呈现逐渐减少的趋势；在实体书阅读时长上，18 岁以后的人群呈现阅读时长逐渐增加的趋势。

2. 从文化程度来看，随着被调查者文化程度的提高，两种阅读时长都大致呈现上升趋势；从阅读数量自评方面来看大致呈正向关系。

3. 从不同的阅读目的来看，大部分人会由于应付学习或工作、了解时事热点和丰富精神世界而增加自己的总阅读时间；当今人们很少会因消遣时间而阅读实体书，但相对的会增加线上阅读的时间，这之间的差异很有可能是由于实体书以及线上阅读的内容不同所导致的。

4. 从阅读满意度方面来看，人们对于自己阅读效率的满意度的提升将会促进阅读时长的增加，这是一个良性循环：对阅读满意度越高，阅读热情也就更盛。

(四) 影响阅读质量满意度的 Logit 定序回归

通过初步筛选挑选出性别、年龄、文化程度、月收入、七种阅读目的及纸质

阅读与总阅读时长比这 12 个解释变量分析其对被调查者阅读质量满意度的影响。为简化模型，满意度简化为：1 表示不满意，2 表示一般，3 表示满意：

表 15 影响阅读质量满意度因素显著性检验表

	LR Chisq	Df	P 值	变量显著性
年龄	16.785	4	0.002	**
性别	2.800	1	0.094	.
文化程度	10.859	4	0.028	*
月收入	39.923	1	0.000	***
应付问题	17.579	1	0.000	***
了解时事	2.351	1	0.125	
增长见识	8.341	1	0.004	**
丰富	6.257	1	0.012	*
消遣时间	4.314	1	0.038	*
个人兴趣	2.131	1	0.144	
其他	10.733	1	0.001	**
纸质阅读与总阅读时长比	69.478	1	0.000	***

通过对各个因素进行显著性分析发现，除阅读目的中的了解时事和个人兴趣外，其余 10 个因素均对阅读质量满意度有显著影响。

为建立一个尽量简单又较为准确的模型，运用 AIC 的模型选择准则，最终选择模型如下：

表 16 logistic 回归系数表

变量	估计值	标准差	P 值
年龄 19 到 24(X ₁)	0.319	0.489	0.251
年龄 25 到 40(X ₁)	-0.594	0.537	0.262
年龄 41 到 50(X ₁)	0.081	0.520	0.376
年龄 50 及以上(X ₁)	0.776	0.572	0.095
性别女(X ₂)	0.301	0.180	0.092
学历高中或中专(X ₃)	-0.275	0.556	0.621
学历大专(X ₃)	-1.243	0.530	0.015
学历大学本科(X ₃)	-1.070	0.524	0.041
学历研究生及以上(X ₃)	-1.214	0.557	0.029
月收入(X ₄)	0.442	0.071	0.000
应付问题(X ₅)	-0.907	0.216	0.000
了解时事(X ₆)	0.300	0.196	0.051
增长见识(X ₇)	0.563	0.196	0.000
丰富精神世界(X ₈)	0.547	0.220	0.013
消遣时间(X ₉)	-0.407	0.197	0.038
个人兴趣(X ₁₀)	0.314	0.215	0.140
其他(X ₁₁)	-1.678	0.513	0.000
纸质与总阅读时长比(X ₁₂)	2.901	0.364	0.000

表 17 logistic 回归截距表

	估计值	标准差	P 值
1 2	0.826	0.713	0.067
2 3	3.871	0.738	0.156

可知模型以年龄为 18 岁及以下的男性，文化程度为初中及以下的被调查者为基准水平建立模型，得到以下结论：

1.相较于男生，女生的阅读满意度通常会更高。

2.对比不同文化程度的被调查者对自身阅读质量满意度发现，随着文化程度的提高，被调查者对自身阅读的满意程度有一定的下降趋势，可能由于人们日益增长的物质文化需要同社会文化不平衡发展之间有冲突和矛盾。

3.对比不同阅读目的，以应付学习或工作中的问题和消遣时间为目的的阅读会降低读者对阅读质量的满意度；以了解时事热点、提高个人兴趣、丰富精神世界和个人兴趣为目的的阅读会较大程度上提高读者对于阅读质量的满意度，反映出被调查者倾向于承认“关注社会、个人发展”为对自己正能量的输入这一观点。

4.纸质阅读时长与总阅读时长之比作为解释变量，能够显著反映出读者阅读质量的满意度，说明纸质阅读时间相较总阅读时长越长，读者对自己的阅读质量满意度会提高。可能相较线上阅读，纸质读物不适应碎片化阅读形式，而较长时间的阅读会给读者带来更多精神享受和深度思考；读者对纸质读物的选取，在质量上也会比线上阅读更有保障。

通过以上分析以及运算，得到各个变量的效应值，可以建立预测读者阅读满意度的模型如下所示，该模型为最小 AIC 下的模型：

$$P(\text{sati} \leq 1) = \frac{\exp(\alpha_1 + [\beta_j, x_j])}{1 + \exp(\alpha_1 + [\beta_j, x_j])}$$

$$P(\text{sati} \leq 2) = \frac{\exp(\alpha_2 + [\beta_j, x_j])}{1 + \exp(\alpha_2 + [\beta_j, x_j])}$$

$$P(\text{sati} = 2) = P(\text{sati} \leq 2) - P(\text{sati} \leq 1)$$

其中， $\alpha_i(i=1,2)$ 表示模型的截距， $\beta_j(j=1,2,3,\dots,12)$ 表示各个变量对应的参数估计； $x_1, x_2, x_3, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}$ 均为二分类变量，即 0, 1 变量， x_4 与 x_{12} 为连续型变量。

（五） 公共阅读空间现状分析

1. 居民前往阅读空间的目的分析

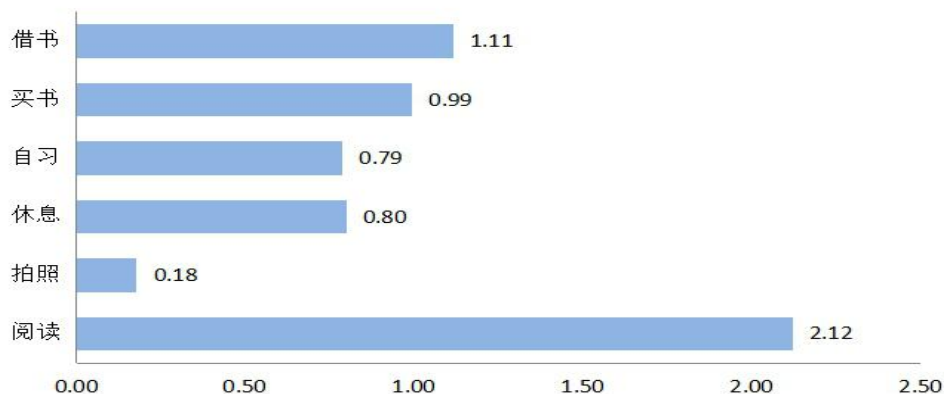


图 9 前去阅读空间目的得分图

由图可知被调查者去阅读空间的主要目的为阅读；借书与买书次之，很少拍照为目的。

2. 阅读空间推广及普及

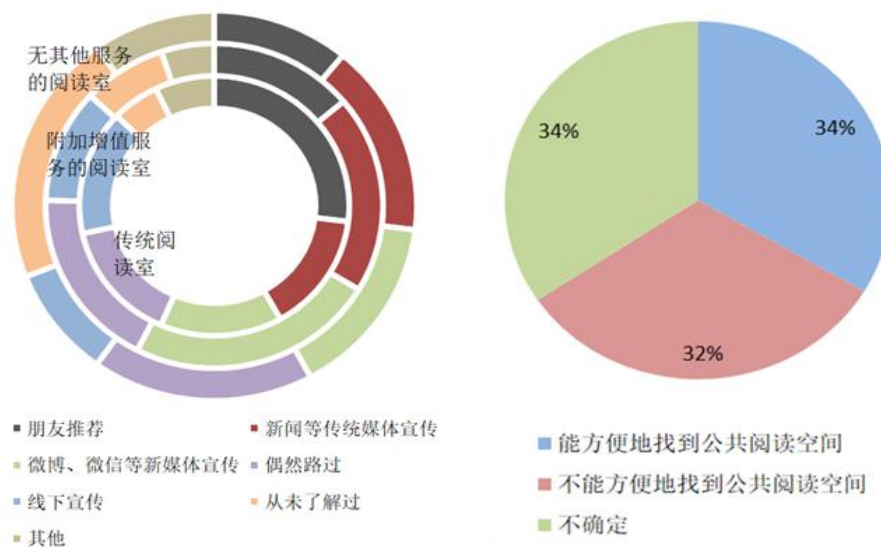


图 10 公共阅读空间推广及普及情况图

由图左图显示传统阅读室普及程度广；附加增值服务的阅读室线下宣传力度较弱，无其他服务的阅读室中宣传效果不明显。图 11 的右图显示能方便地找、不能方便地找、不确定能方便地找到公共阅读空间的比例分布均匀。

3. 读者偏好与阅读空间优势

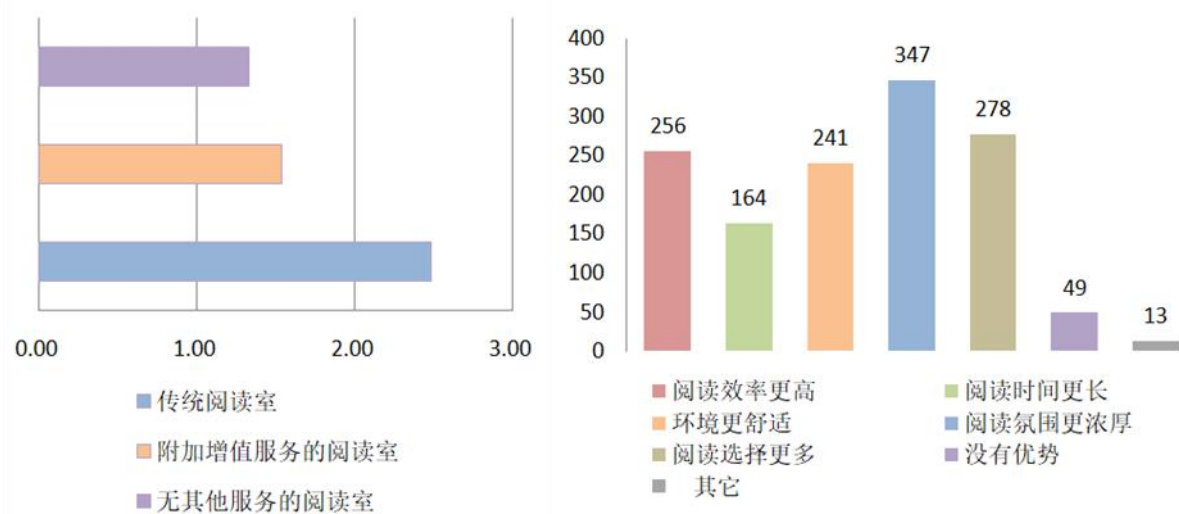


图 11 读者偏好与阅读空间优势图

由左图发现，被调查者对传统阅读室的偏好程度较大，对附加增值服务的阅读室的偏好程度略高于无其他服务的阅读室。

由右图可知，相较于在私人阅读空间，大多数被调查者认为在公共阅读空间阅读有很多优势：其中阅读氛围更浓厚、阅读选择更多、阅读效率更高和环境更舒适这四个方面最为突出。

4. 不同类型阅读空间优劣势分析

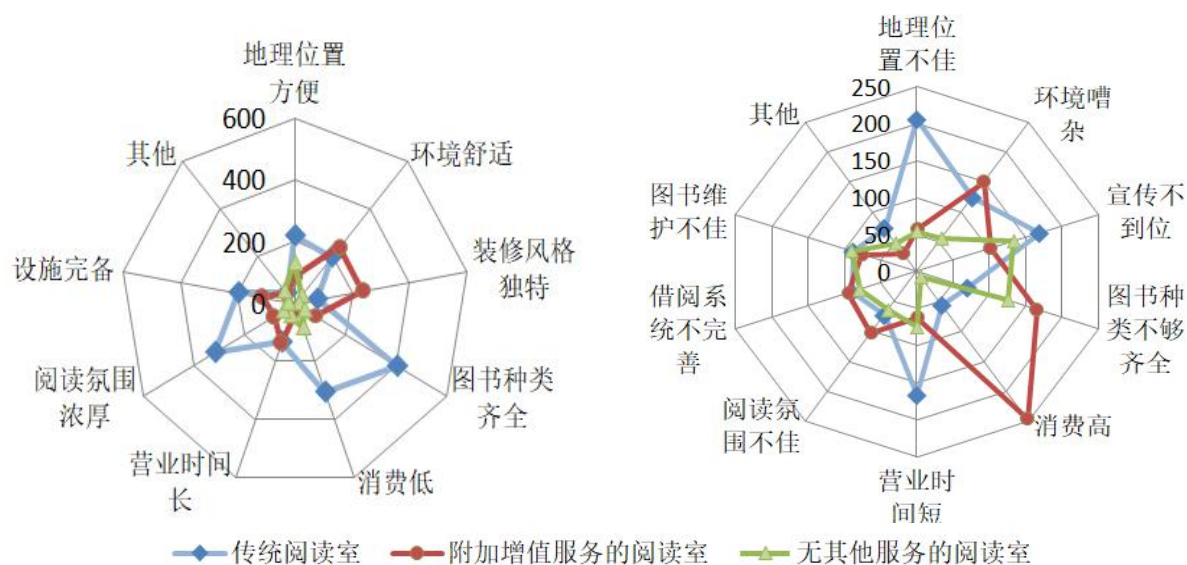


图 12 不同类型阅读空间优劣势图

由左图对比不同类型阅读空间的优势发现，传统阅读室除装修风格与环境舒适度外的其他优势都很突出。附加增值服务的阅读室在装修风格与环境舒适度两个指标上有明显优势，是吸引读者、增强竞争优势的强有力因素。无其他服务的阅读室在所有指标上都不是很突出。

由右图对比不同类型阅读空间的不足发现，传统阅读室地理位置不佳、宣传不到位、营业时间短是其最突出的问题；附加增值服务的阅读室环境嘈杂、消费高，这与其经营模式，及该类图书馆还为使用者提供更多元的活动和服务密切相关；无其他服务的阅读室自身宣传不到位、图书种类也不够齐全。

六、 公共阅读空间需求影响因素分析

为进一步了解当代人对公共阅读空间不同因素的需求情况,选择主成分法对样本数据进行探索性因子分析,在提取公因子后分别计算各指标重要性程度均值,通过均值比较人们对公共阅读空间不同因素的需求情况;接下来结合前文描述统计,采用 Apriori 关联规则挖掘算法,为三种因子所指向的三种不同类型阅读空间定位目标客户群体;最后结合验证性因子分析,建立 PLS 路径模型,探究各因子的直接与间接影响,为公共阅读空间的建设和提出建议。

(一) 阅读空间需求因素的探索性因子分析——主成分法

运用 SPSS20.0 软件对统计而得的样本数据进行探索性因子分析。为区别不同的三级指标对总体满意度的影响程度大小,筛选出对总指标有显著影响的因子。在提取初始因子时采取主成分法。初始因子载荷矩阵由于因子在原变量上的载荷值相差不大,故不利于解释它们的含义,因此需要进行因子旋转,这里选择最大方差法进行分析。

表 18 方差贡献率

成份	旋转平方和载入		
	合计	方差的 %	累积 %
1	7.624	38.120	38.120
2	3.634	18.171	56.291
3	2.213	11.063	67.354
4	1.476	7.379	74.733

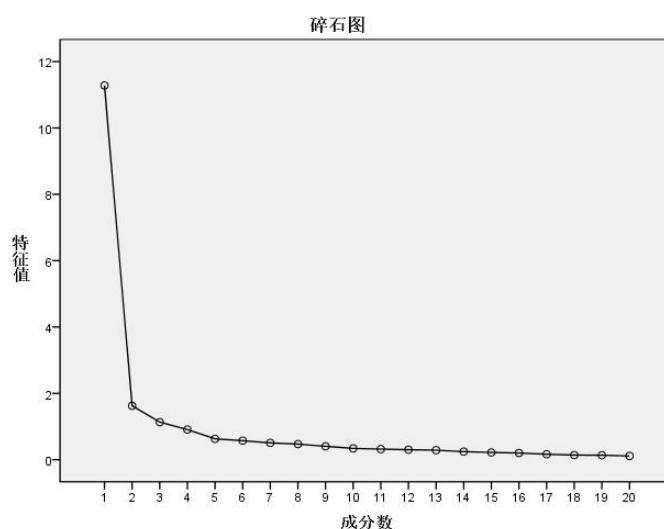


图 13 碎石图

由图和表可知，提取 4 个因子时，共解释了原有变量总方差的 74.733%，总体上，原有变量的信息丢失较少，因子分析的效果较理想。

表 19 旋转成份矩阵

指标 ^a	成份 ^a				命名 ^a
	1 ^a	2 ^a	3 ^a	4 ^a	
环境卫生 ^a	.865 ^a	.115 ^a	.111 ^a	.086 ^a	基础环境 ^a
阅读秩序维护 ^a	.839 ^a	.310 ^a	.143 ^a	.018 ^a	
环境安静 ^a	.806 ^a	.344 ^a	.131 ^a	-.032 ^a	
设施质量 ^a	.782 ^a	.070 ^a	.042 ^a	.338 ^a	
图书种类齐全 ^a	.762 ^a	.370 ^a	.178 ^a	.011 ^a	
交通便利 ^a	.748 ^a	.181 ^a	.062 ^a	.137 ^a	
图书维护 ^a	.713 ^a	.323 ^a	.363 ^a	.053 ^a	
营业时间长 ^a	.702 ^a	.413 ^a	.028 ^a	.257 ^a	
标识清晰易懂 ^a	.663 ^a	.508 ^a	.296 ^a	.030 ^a	
服务态度良好 ^a	.652 ^a	.435 ^a	.250 ^a	.168 ^a	
场地宽敞 ^a	.644 ^a	.597 ^a	.164 ^a	.099 ^a	
阅读氛围浓厚 ^a	.620 ^a	.391 ^a	.417 ^a	-.030 ^a	
便民设施完备 ^a	.509 ^a	.371 ^a	.368 ^a	.228 ^a	
融合线上服务 ^a	.327 ^a	.761 ^a	.286 ^a	.076 ^a	惠民便捷 ^a
自助服务水平 ^a	.461 ^a	.755 ^a	.201 ^a	.120 ^a	
经济实惠 ^a	.573 ^a	.606 ^a	.114 ^a	.123 ^a	
定期开展阅读活动 ^a	.210 ^a	.044 ^a	.850 ^a	.187 ^a	拓展阅读 ^a
提供亲子阅读室 ^a	.027 ^a	.368 ^a	.768 ^a	.090 ^a	
提供增值服务 ^a	-.090 ^a	.438 ^a	.122 ^a	.759 ^a	附加价值 ^a
装修风格特色 ^a	.385 ^a	-.132 ^a	.213 ^a	.721 ^a	

由上表可以发现读者对公共阅读空间的需求情况基本可以概括成以下四点：基础环境因子、惠民便捷因子、拓展阅读因子以及附加价值因子。

为了比较各因子的重要程度，分别计算四个因子所含的原有变量的重要性均值，并对其进行比较，计算结果如下图所示。

表 20 因子载荷矩阵

指标	因子载荷	均值	命名
环境卫生	.865	3.987	基础环境 (均值 3.933)
阅读秩序维护	.839	4.164	
环境安静	.806	4.260	
设施质量	.782	3.733	
图书种类齐全	.762	4.061	
交通便利	.748	3.755	
图书维护	.713	3.980	
营业时间长	.702	3.797	
标识清晰易懂	.663	3.866	
服务态度良好	.652	3.971	
场地宽敞	.644	3.893	
阅读氛围浓厚	.620	3.974	
便民设施完备	.509	3.685	
融合线上服务	.761	3.637	惠民便捷 (均值 3.735)
自助服务水平	.755	3.713	
经济实惠	.606	3.855	
定期开展阅读活动	.850	3.241	拓展阅读 (均值 3.217)
提供亲子阅读室	.768	3.193	
提供增值服务	.759	3.011	附加价值 (均值 3.037)
装修风格特色	.721	3.063	

从重要程度均值来看，基础环境因子是人们认为最重要的因子。其中阅读秩序维护与环境安静的因子载荷与重要性程度均位于前列，因此，加强硬件设施的配置与维护、营造良好的阅读环境是所有公共阅读空间必须施行的举措。

惠民便捷因子的重要程度仅次于基础环境因子。在“互联网+”不断推进的当今社会，便捷的线上服务与自助服务成为人们关注的焦点。随着餐饮、文创等业态融入，越来越多的人希望公共阅读空间在价格上足够亲民。

拓展阅读因子下的原有变量重要程度均值普遍低于前两个因子，总均值为 3.217。重要程度均值最低的是附加价值因子。由此可见，人们前往公共阅读空间的主要目的仍是阅读，而不是享受其带来的附加价值。

（二） 阅读空间需求因素验证性因子分析——PLS 路径模型

本项目中依据前文探索性因子分析降维结果，得到公共阅读空间影响因素 PLS 路径模型包含 4 个潜变量：基础环境、拓展阅读、附加价值、惠民便捷以及量表中 20 个显变量。

使用 R 软件画出模型路径图，建立公共阅读空间影响因素模型如下：

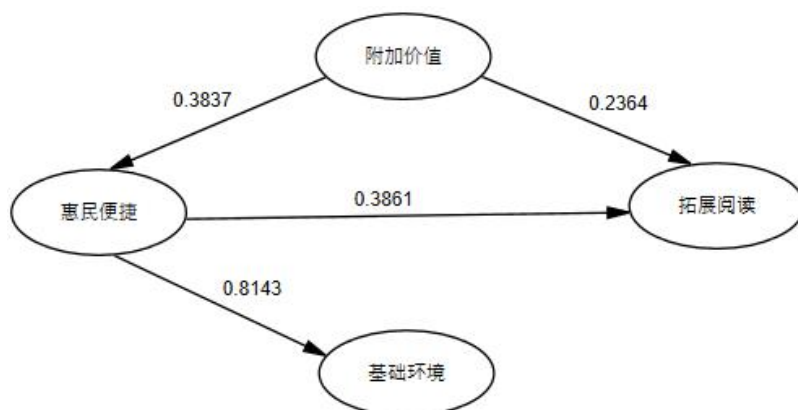


图 14PLS 路径模型图

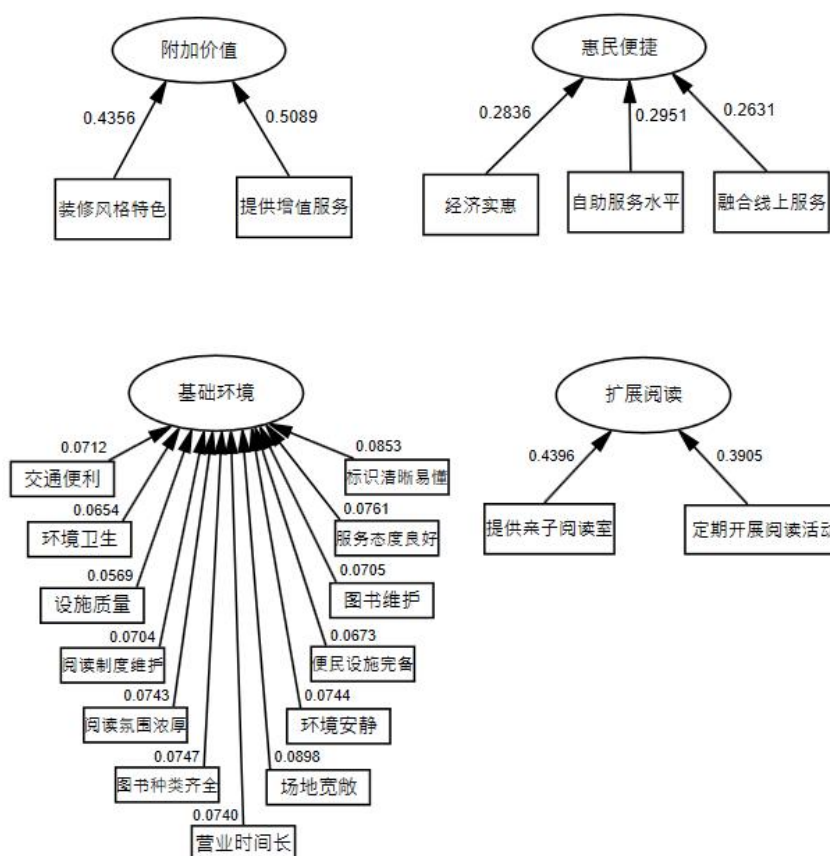


图 15PLS 路径模型权重系数图

表 21 测量模型检验表

	C.alpha	DG.rho
附加价值	0.4621	0.7880
惠民便捷	0.8914	0.9328
基础环境	0.9610	0.9655
拓展阅读	0.7077	0.8725

由表中的 Cronbach's alpha 检验是否具有内部一致性，其中附加价值 alpha

值为 0.4621，属于中信度范围，其余因子 alpha 值均大于 0.7，属于高信度范围。通过 DG.rho 值检验复合信度，发现其值均达到 0.7 以上，测量工具可靠。

表 22 结构模型检验表

	R^2	共同度	冗余度	AVE
附加价值	0.0000	0.6490	0.0000	0.6490
惠民便捷	0.1472	0.8217	0.1210	0.8217
基础环境	0.6631	0.6835	0.4532	0.6835
拓展阅读	0.2749	0.7725	0.2124	0.7725

由表可知，各因子共同度均在 0.60 以上，AVE 值处于 0.6490~0.8217 之间，均大于 0.6，表明本项目所使用的量表具有较好的效度。整体模型的拟合优度为 0.5067，拟合效果可以接受。

表 23PLS 路径模型效应表

relationships	direct	indirect	total
附加价值 ->惠民便捷	0.384	0.000	0.384
附加价值 ->基础环境	0.000	0.312	0.312
附加价值 ->拓展阅读	0.236	0.148	0.384
惠民便捷 ->基础环境	0.814	0.000	0.814
惠民便捷 ->拓展阅读	0.386	0.000	0.386
基础环境 ->拓展阅读	0.000	0.000	0.000

由表可得，惠民便捷因子与基础环境因子的关系最为密切，其次为惠民实用对拓展阅读和附加价值对惠民实用，附加价值与拓展阅读存在一定的正相关性但相关性不强。附加价值对基础环境存在间接影响 $\beta = 0.384 \times 0.814 = 0.312$ 。

联系因子分析结果发现，基础环境因子是人们认为最重要的因子，对公共阅读空间基础环境的改进更贴合读者的切实需求，能够促使民众主动进行阅读。结合 PLS 路径模型得知，公共阅读空间可以通过改善惠民便捷因子（例如增加线上服务、提供自助借阅机等方式）达到改进基础环境的目的，且这种方式是最直接有效的，通过对惠民便捷因子的改善可以同时改进拓展阅读因子；也可以通过改善附加价值因子（例如融合餐饮、文创等多元化业态）间接带动基础环境的改善，为人们提供良好的阅读环境，提供更为优质的现代公共文化服务。

（三） 各类公共阅读空间目标群体定位——关联规则

为了探究不同类型阅读室的目标客户群体，本项目采用 Apriori 关联规则挖

掘算法，分别探究认为惠民便捷因子、拓展阅读因子以及附加价值因子重要的人群特征，为三种因子所指向的三种不同类型阅读室定位目标客户群体。

1. 模型构建

本项目选择探究前项为基本信息及个人阅读现状，后项为各因子的关联规则，选择 Apriori 算法挖掘满足该条件的关联规则并提取后项为“惠民便捷=重要”且置信度最高的前五条规则进行分析。首先对前项进行如下分类处理：

表 24 前项数据预处理表

前项	选项	分类
文化程度	初中及以下、高中或中专	低
	大专、大学本科、研究生及以上	高
月收入	3000 元及以下	低
	3001~8000 元	中等
	8001 元以上	高
每周阅读总时长、实体书阅读总时长	前 33%	少
	33%~67%	中等
	67%~100%	多
阅读数量自评	非常少、较少	少
	一般	一般
	非常多、较多	多
阅读质量满意度	非常不满意、不太满意	不满意
	一般满意	一般
	非常满意、比较满意	满意

在对后项进行预处理时，将因子得分小于均值划分为该因子不重要，大于均值为该因子重要，设定最小支持度为 0.1，最小置信度为 0.5。

2. 关联规则结果分析

使用 R 语言进行关联规则挖掘得到各因子的关联规则如下所示：

表 25 惠民便捷因子关联分析表

前项	后项	支持度	置信度	提升
性别=男 文化程度=高 增长见识=否 最常用阅读媒介=手机 阅读质量满意度=一般	惠民便捷=重要 → 无其他服务的阅 读室	0.1015	0.8462	1.7707
性别=男 文化程度=高 增长见识=否 最常用阅读媒介=手机		0.1181	0.8312	1.7394
文化程度=高 增长见识=否 最常用阅读媒介=手机 每周阅读总时长=多 阅读质量满意度=一般		0.1089	0.8310	1.7390
文化程度=高 增长见识=否 最常用阅读媒介=手机 每周阅读总时长=多		0.1125	0.8243	1.7250
性别=男 文化程度=高 增长见识=否 每周阅读总时长=多		0.1033	0.8235	1.7234

由上表可以看出，第二条规则是第一条的子集，提升度表明满足前项条件的人群更重视惠民便捷因子的概率分别提升 0.7394 倍和 0.7707 倍。即阅读质量满意度一般、最常用手机进行阅读，阅读主要目的不是增长见识的高学历男性更重视惠民便捷因子，因而此类人群更倾向于选择该因子指向的无其他服务的阅读室。

同理第四条规则是第三条规则的子集，表明阅读质量满意度一般、最常使用手机阅读、阅读时长多、阅读主要目的不是增长见识的高学历人群更重视惠民便捷因子。第五条规则表明阅读主要目的不是增长见识、每周有较多的阅读时长的的高学历男性更重视惠民便捷因子。

由此可见，无其他服务的阅读室目标客户主要集中在阅读时长较多的高学历群体，该类群体有一定的阅读基础与较好的阅读习惯，更重视阅读过程的自主性与便捷性，与阅读室提供的增值服务相比，经济上实惠亲民显得更为重要。

表 26 拓展阅读因子关联分析表

前项	后项	支持度	置信度	提升
增长见识=否 丰富精神世界=否 最常用阅读媒介=手机 基础环境=重要	拓展阅读=重要 → 传统阅读室	0.1107	0.9524	1.8112
增长见识=否 丰富精神世界=否 阅读数量自评=少 基础环境=重要		0.1070	0.9206	1.7508
增长见识=否 丰富精神世界=否 发展个人兴趣=否 阅读数量自评=少 基础环境=重要		0.1015	0.9167	1.7433
每周实体书阅读总时长=中等 阅读数量自评=少 基础环境=重要		0.1070	0.9063	1.7235
增长见识=否 丰富精神世界=否 基础环境=重要		0.1421	0.9059	1.7228

由上表第 1、2、3、5 条规则可知，认为公共阅读空间基础环境重要、阅读主要目的不是增长见识、丰富精神世界的群体，与样本总体相比更重视拓展阅读因子。由前四条规则可知，自评阅读数量少、经常使用手机进行阅读且每周实体书阅读总时长中等的人群更重视拓展阅读因子。五条规则的提升均大于 1.7。

从中可以发现，拓展阅读因子指向的传统阅读室目标客户群体定位在阅读量较少、以线上阅读为主的人群，该类人群希望通过阅读室提供相关的拓展性阅读活动达到促进阅读的目的。

五条规则都反映出重视拓展阅读的人群普遍看重公共阅读空间的基础环境，因而传统阅读室应在做好基础环境建设的前提下进一步开展拓展阅读建设。

表 27 附加价值因子关联分析表

前项	后项	支持度	置信度	提升
增长见识=否 丰富精神世界=是 惠民便捷=重要	附加价值=重要 → 附加增值服务的 阅读室	0.1089	0.9516	1.9390
文化程度=高 增长见识=否 丰富精神世界=是 惠民便捷=重要		0.1070	0.9508	1.9374
年龄=41~50岁 丰富精神世界=是 消遣时间=否 惠民便捷=重要		0.1033	0.8889	1.8112
阅读数量自评=一般 基础环境=不重要 惠民便捷=重要		0.1015	0.8871	1.8075
文化程度=高 阅读数量自评=一般 基础环境=不重要 惠民便捷=重要		0.1015	0.8871	1.8075

从上表可以看出第一条规则是第二条规则的子集，反映出阅读目的为丰富精神世界、重视惠民便捷因子的高学历人群更重视公共阅读空间的附加价值。第四条规则是第五条规则的子集，即自评阅读数量一般、认为公共阅读空间基础环境不重要而惠民便捷性重要的高学历人群更看重附加价值因子。第三条规则反映出阅读目的为丰富精神世界而非消遣时间、认为惠民便捷因子重要的41~50岁人群更看重附加价值因子。

通过上述分析发现，附加价值指向的附加增值服务的阅读室（例如网红书店、图书馆餐厅等）目标客户群体为阅读数量一般，希望通过阅读提升文化修养的高学历中年人群，该类人群对阅读空间的基础环境例如相关设施、藏书量、阅读氛围等没有过多追求，他们更重视阅读空间的实用性与多元化布局，享受阅读的情调。

七、 结论与建议

（一） 结论

1. 前往阅读空间目的的多样化，文化素养与阅读需求成正相关关系

公共阅读空间的主要顾客来源于具有较高学历的读者。但同时，具有较高文

化基础的读者相较于其他读者，对于自己阅读现状的满意度更低，文化程度无法与其阅读质量相匹配，也表明人们日益增长的物质文化需要同社会文化不平衡发展之间有冲突和矛盾。

2. 电子阅读媒介普及广泛，存在的问题凸显

被调查的读者中，大多数都会通过手机进行阅读，其比例远远大于购买纸质书和节约纸制书。读者虽然习惯于通过手机进行较为碎片化的浅阅读，但这并不能给予其精神世界的充实感和满足感。

3. 公共阅读空间具有促进阅读的作用，但普及程度欠佳

提高阅读氛围最为简单、高效的方式便是在一个小范围空间内为读者提供远离外界纷扰、藏书齐全的阅读空间。虽然传统阅读室和附加增值服务的商业化阅读室通常为大众所熟知，但无其他服务的阅读室，却缺乏宣传。

4. 不同阅读空间的目标群体不同，存在不同的局限性

传统阅读空间的主要局限性在于资源过于集中化，导致其虽然空间大、馆藏丰富，但并不日常化，其目标群体普遍阅读量较少、以线上阅读为主。无其他服务的阅读室为满足公众的阅读日常化，但常常因为经费不足、书籍过少，无人知晓等原因无法将开设目的落到实处，其目标群体主要为阅读时间较长的高学历群体。附加增值服务的阅读空间通过其餐饮服务、独特风格吸引大量读者，但阅读和商业化的平衡点难以把握，此类空间更加吸引阅读数量一般但希望追求更深层次文化提升的高学历中年人群。

5. 环境因素显著影响阅读空间价值，应当予以重视

从因子得分结果来看，环境因素对于公共阅读空间的影响最为显著，阅读环境不完善将会影响到阅读空间价值的体现和读者的阅读感受，问题严重将会导致读者对公共阅读空间的不满。

(二) 建议

1. 宏观层面关于战略文化布局的建议

(1) 城市公共阅读空间的规划布局，其理论的逻辑起点在于文化空间。

(2) 注重宏观层面的战略文化空间设计，政府要有相应的政策导向，应对城市阅读空间的类型、布局、建设进行整体规划。同时，地方政府应该制定相关政策，引导资本和社会力量从事阅读空间建设。

(3) 加强对不同阅读空间行业的监管和引导，合理空间分布。

2. 中观层面关于文化空间完善的建议

(1) 空间结构优化

城市公共阅读空间作为一个有机的整体,应保证在布局上的公共性和便利性,各类型阅读空间的资源配置与合作布局,避免信息资源的重复建设和浪费。

(2) 公共阅读空间的改善

1. 公共图书馆作为国家全公益和零门槛的阅读空间应最先实现文化资源的下沉,以社区书屋和独立书店为布局点进行总分馆的实践,实现文化资源的转借功能,增强阅读服务便利性和文化资源的利用率。

2. 独立书店可以依托社区布局“社区+书店”的差异化、便民化服务模式,促进不同公共阅读设施之间的弹性使用,实现不同类型阅读空间之间以资源配置和便民服务为前提的积极融合。

3. 各类型阅读空间需布局虚拟阅读空间,形成以线上图书馆为依托的文化流动空间,依靠 O2O 平台实现线上转借线下空间以增强体验,实现空间服务的转型。

4. 通过多样化渠道向潜在目标读者的群体进行宣传。

3. 微观层面关于阅读氛围营造的建议

(1) 提升阅读空间服务质量,从而提升阅读氛围。阅读空间环境应做到整洁、安静,提供必要的残疾人通道,为需要的人群配备老花镜、放大镜等设备。将阅览书籍与售卖书籍进行区分,减少不必要的图书损耗。根据实际情况调整经营时段以充裕居民读书时间。

(2) 提高公众阅读兴趣、热情。举办线下的读书沙龙、阅读分享会、作者见面会等活动,进一步提高公众对阅读的需求和体验。

4. 各类公共阅读空间建设框架

根据不同类型的公共阅读空间发展现状提出相应的建设构想，如下图所示：

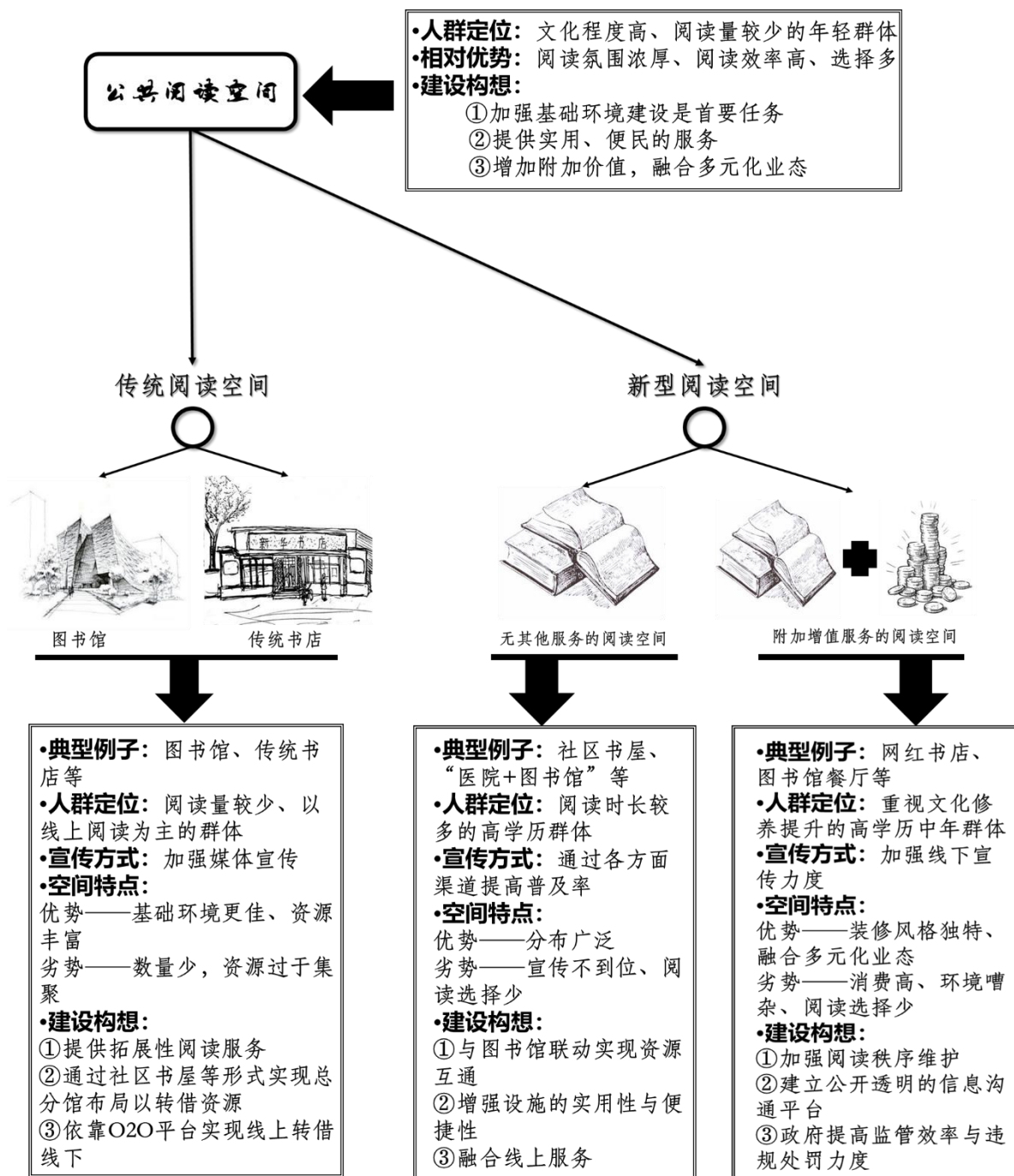


图 16 公共阅读空间建设框架图

参考文献

- [1]王炎龙,郭玉.基于文化规划视角的城市公共阅读空间多维布局探究[J].中国出版,2018(18):3-8.
- [2]叶帆.公共图书馆打造“城市书房”案例研究[D].安徽大学,2018.
- [3]刘雅薇.我国独立书店的发展与转型[D].华中科技大学,2015.
- [4]侯凤芝.城乡融合发展中农村公共阅读空间建构模式研究[J].中国出版,2018(09):29-33.
- [5]刘艳.基于“SoLoMo+O2O”模式的城市地下流动阅读空间构建策略[J].国家图书馆学刊,2017,26(04):22-28.
- [6]胡卉,李睿.面向文化成都建设的图书馆阅读推广策略构建[J].图书馆理论与实践,2015(06):75-79.
- [7]刘润,杨永春,任晓蕾.1990s 末以来成都市文化空间的变化特征及其驱动机制[J].经济地理,2017,37(02):114-123.
- [8]申立,陆巍,王彬.面向全球城市的上海文化空间规划编制的思考[J].城市规划学刊,2016(03):63-70.
- [9]万宇.多元理论维度中的儿童阅读空间构建——以我国公共图书馆儿童阅览室为例[J].图书馆杂志,2014,33(04):42-47.
- [10]王子舟.我国公共阅读空间的兴起与发展[J].图书情报知识,2017(02):4-12.
- [11]伍乐平,张晓萍.国内外“文化空间”研究的多维视角[J].西南民族大学学报(人文社科版),2016,37(03):7-12.
- [12]蒋宏宇,李理.公共体育服务多元供给中的政府责任及其实现路径[J].湖南科技大学学报(社会科学版),2018,21(04):165-171.
- [13]夏立新,郑路,代沁泉,毕崇武.全民阅读语境下的个人阅读共享空间构建研究[J].情报科学,2018,36(02):9-13.
- [14]赵富强.基于 PLS 路径模型的顾客满意度测评研究[D].天津大学,2010.
- [15]李实,叶强,李一军,罗嗣卿.挖掘中文网络客户评论的产品特征及情感倾向[J].计算机应用研究,2010,27(08):3016-3019.
- [16]赵松. Apriori 算法的改进及应用[D].哈尔滨理工大学,2006.
- [17]曾孝文.关联规则数据挖掘方法的研究[J].计算机与现代化,2006(09):90-92+96.
- [18]薛薇.统计分析与 SPSS 的应用[M].中国人民大学出版社,2014.
- [19]李金昌,苏为华.统计学[M].机械工业出版社,2014.

[20]吴喜之.高等院校研究生用书复杂数据统计方法基于 R 的应用[M].北京:中国人民大学出版社.2015.

[21]北晚新视觉网.国民阅读调查报告发布,能读这么多书的都是“万中无一”[EB/OL].

[2018-04-19].<http://baijiahao.baidu.com/s?id=1598138775977982752&wfr=spider&for=pc>

附录一：调查问卷

公共阅读空间现状及前景调查

问卷编号：访问员：访问地点：

尊敬的朋友：

您好！我们是XX大学的学生，目前正在做一项关于“公共阅读空间现状及前景”的调查，希望通过本次调查，了解当前杭州市居民的阅读现状与公共阅读空间的发展现状。我们会对您的回答严格保密。非常希望您能完成此份问卷，谢谢！

一、基本信息

1.您的性别

- (1) 男 (2) 女

2.您的年龄

- (1) 18岁及以下 (2) 19~24岁 (3) 25~40岁
(4) 41~50岁 (5) 51~60岁 (6) 60岁以上

3. 您的文化程度

- (1) 初中及以下 (2) 高中或中专 (3) 大专
(4) 大学本科 (5) 研究生及以上

4.您的职业

- (1) 学生 (2) 企业管理者 (3) 企业一般工作人员
(4) 机关事业单位领导干部 (5) 机关事业单位普通工作人员 (6) 教师
(7) 医生 (8) 服务人员 (9) 军警 (10) 个体户 (11) 自由职业者
(12) 农民 (13) 工人 (14) 农民工 (15) 离退休人员
(16) 下岗失业人员 (17) 无业人员 (18) 其他_____

5.您的月收入

- (1) 无收入 (2) 3000元及以下 (3) 3001~5000元
(4) 5001~8000元 (5) 8001~10000元 (6) 10000元以上

二、居民阅读现状

6.您阅读的目的是（多选）

- (1) 应付学习或工作中的问题 (2) 了解时事热点 (3) 增长见识、拓展思维
(4) 丰富精神世界、提升文化修养 (5) 消遣时间 (6) 发展个人兴趣
(7) 其他_____

7.您所喜欢的阅读类型有（多选）

- (1) 哲学类 (2) 社会科学类 (3) 自然科学类 (4) 文学类 (5) 教育类
(6) 生活类 (7) 新闻热点 (8) 杂志期刊 (9) 漫画图集 (10) 其他

8.您通过什么媒介进行阅读

- (请根据使用频率从高到低排序) __、__、__、__、__
(1) 借阅纸质书 (2) 购买纸质书 (3) 手机 (4) 电子阅读器 (kindle 等) (5) 电脑

9.您每周阅读（线上+线下）总时长为_____小时

10.您每周实体书阅读总时长为_____小时

11.您对自己阅读数量的评价是

- (1) 非常少 (2) 较少 (3) 一般 (跳过第 12 题)
(4) 较多 (跳过第 12 题) (5) 非常多 (跳过第 12 题) (6) 不清楚
(跳过第 12 题)

12.阅读数量少的原因（多选）

- (1) 阅读时间少 (2) 缺乏阅读动力 (3) 缺少合适的阅读环境
(4) 来自外界的干扰因素强 (5) 缺乏感兴趣阅读类型的信息来源 (6)
其他_____

13.您对自己的阅读质量的满意度是

- (1) 非常不满意 (2) 不太满意 (3) 一般 (4) 比较满意 (5) 非常满意

三、公共阅读空间现状

14. 请选出您前往公共阅读空间的三个主要目的（按重要程度排序）

—、—、—

- (1) 阅读 (2) 拍照 (3) 休息 (4) 自习 (5) 买书 (6) 借书

15. 您对不同类型的公共阅读空间的了解途径为

	传统阅读室 (图书馆、 传统书店 等)	附加增值服务的阅 读室 (咖啡厅、图书馆餐 厅、网红书店等)	无其他服务的阅 读室 (城市书房、社区 阅读室等)
朋友推荐			
新闻等传统媒体宣传			
微博、微信等新媒体宣 传			
偶然路过			
线下宣传			
从未了解过			
其他			

16. 您更习惯去什么地方进行纸质阅读

- (1) 传统阅读室（图书馆、传统书店等）
(2) 附加增值服务的阅读室（咖啡厅、图书馆餐厅、网红书店等）
(3) 无其他服务的阅读室（城市书房、社区阅读室等）
(4) 私人场所（家中、办公室等）
(5) 其他_____

17. 您认为在公共阅读空间阅读与个人阅读相比（多选）

- (1) 阅读效率更高 (2) 阅读时间更长 (3) 环境更舒适 (4) 阅读氛
围更浓厚
(5) 阅读选择更多 (6) 没有优势 (7) 其他_____

18.您认为不同公共阅读空间的优势是（多选）

	传统阅读室 (图书馆、传统书店等)	附加增值服务的阅读室 (咖啡厅、图书馆餐厅、网红书店等)	无其他服务的阅读室 (城市书房、社区阅读室等)
地理位置方便			
环境舒适			
装修风格独特			
图书种类齐全			
消费低			
营业时间长			
阅读氛围浓厚			
设施完备			
其他			

19.您认为不同公共阅读空间存在哪些问题（多选）

	传统阅读室 (图书馆、传统书店等)	附加增值服务的阅读室 (咖啡厅、图书馆餐厅、网红书店等)	无其他服务的阅读室 (城市书房、社区阅读室等)
地理位置不佳			
环境嘈杂			
宣传不到位			
图书种类不够齐全			
消费高			
营业时间短			
阅读氛围不佳			
借阅系统不完善			
图书维护不佳			
其他			

20.当您需要一个阅读空间时，在附近能否方便地找到公共阅读空间

- (1) 能
- (2) 否
- (3) 不确定

四、对阅读空间的需求

21. 请根据您对以下公共阅读空间的偏好程度从高到低排序

- (1) 传统阅读室（图书馆、传统书店等）
- (2) 附加增值服务的阅读室（咖啡厅、图书馆餐厅、网红书店等）
- (3) 无其他服务的阅读室（城市书房、社区阅读室等）

以下是公共阅读空间需求量表，请根据您的自身情况，选择您对各指标的需求程度（“1”代表完全不重要，“5”代表非常重要）

因素	需求度				
	完全不重要→非常重要				
交通便利	1	2	3	4	5
环境卫生	1	2	3	4	5
设施质量	1	2	3	4	5
装修风格特色	1	2	3	4	5
阅读秩序维护	1	2	3	4	5
定期开展阅读活动	1	2	3	4	5
提供亲子阅读室	1	2	3	4	5
阅读氛围浓厚	1	2	3	4	5
图书种类齐全	1	2	3	4	5
提供增值服务（如餐饮）	1	2	3	4	5
营业时间长	1	2	3	4	5
环境安静	1	2	3	4	5
经济实惠	1	2	3	4	5
场地宽敞	1	2	3	4	5
便民设施完备	1	2	3	4	5
图书维护	1	2	3	4	5
服务态度良好	1	2	3	4	5
馆内标识清晰易懂	1	2	3	4	5
融合线上服务	1	2	3	4	5
自助服务水平	1	2	3	4	5

附录二：调查照片

