

基于 SPSS 的百度指数与游客量相关性研究

◎ 国 敏

(北京现代职业技术学院 北京 101300)

【摘要】信息化技术的普及及网络平台的建立为旅游研究提供了重要的数据基础。本文以北京故宫为例,基于相关性分析方法,对所选取网络关键词的搜索量指数与实际游客量进行了相关分析。结果表明网络关键词搜索量指数与实际游客量之间存在明显的相关关系;游客一般在出游前 2—5 天对所关注信息进行查询,最关注的出游信息主要为景区门票的价格、相关旅游路线图和景区开放时间。本研究对旅游管理部门做出科学合理的决策具有一定的现实意义,为游客量的预测研究提供了重要的理论依据。

【关键词】SPSS 百度指数 游客量 相关性分析

一、前言

随着我国经济的迅猛发展,旅游业已成为中国经济发展的支柱性产业之一。越来越多的人通过参与旅游活动丰富物质文化生活,因此也带来诸多旅游城市游客量的逐年攀升。尤其一些热门的旅游景点,在“五一”、“十一”等大小长假期间更是出现游客量爆棚甚至超负荷承载的状态。因此,关于游客量的分析研究对城市旅游业和景区生态环境的可持续发展、旅游资源的合理规划和开发具有重要的现实意义。

互联网技术的普及,大数据、云平台等新兴技术的成熟落地,为蓬勃发展的旅游业注入了新鲜的血液。旅游企业选择互联网这一重要平台发布旅游信息,游客的需求和体验也不再局限于传统的旅游方式,互联网技术可随时为游客提供个性化的服务和实时有效的旅游信息。越来越多的旅游者通过网络搜索技术,有针对性地获取实时有效的旅游信息。根据 CNNIC 的研究显示,搜索引擎是旅游用户在线获取旅游信息的首选渠道,占比近 8 成。

可见,旅游业与信息化的融合愈加突显。运用大数据、搜索引擎等互联网技术分析旅游的网络关注度与现实旅游行为的关系,对旅游管理部门做出科学、合理的决策和安排具有重要的现实意义,为游客量的预测研究提供重要的理论依据。

二、国内外相关研究

目前,基于 SPSS 的游客量相关研究的数据基础主要

以实际游客量的历史数据为主。张红贤等收集了 1978—2003 年中国入境游客量的值,运用多元回归分析的方法,对我国入境游客的规模进行了预测。黄宏等基于 1979—2007 年中国入境旅游人数资料,利用 SPSS 统计分析软件,预测中国入境旅游人数。结果表明,该方法准确可靠高、方便快捷,适用于入境游客量的预测。李世泰等基于山东省 1992—2008 年国内游客量,利用 SPSS 软件定量分析模型分析了不同时间尺度旅游客源市场预测方面的客观差异。随着基于信息技术对社会行为的深入研究,网络搜索技术的旅游行为预测研究主要集中在网络关注度与旅游客流关系的研究。路紫、赵亚红、吴士锋、韩冰运用网上查询系统、网站访问数量统计工具,获取旅游网站访问者的时间分布资料,分析了网络信息流对旅游流的导引作用。梁明英、王丽娜以山东泰山为例,通过对景区游客进行问卷调查,调研了旅游网络信息在景区游客中的应用,分析出景区旅游网络信息对该旅游目的地游客行为模式的影响。戴维森等 (Davidson, et al.) 通过实证对比分析以台湾为旅游目的地的旅游网站的信息流与现实旅游流,发现网站信息流对游客旅游行为具有引导潜力。

以上国内外文献研究显示,以往研究主要基于实际游客量的历史数据的 SPSS 相关性分析和网络搜索技术两种方法分别对游客量进行了论证分析,但将两种方法相结合的分析研究文献较少。本文将借助 SPSS 软件,通过收集能够反映旅游关注度的网络关键词的百度搜索量指数,对北京故宫在 2010 年五个假期的实际游客量进行相关性分析。

三、相关性分析

1、关键词的选取

由于游客的异质性,他们对事物关注的角度会有差异,因此在出游前检索的相关信息和网络关键词具有多样性。根据旅游理论基础、游客的行为特点以及游客在出游前对旅游目的地信息的关注,对游客在出游前对旅游目的地(景区景点)旅游信息所关注的内容的分析,选取出旅游目的地(景区景点)的名称、图片、旅游路线、门票价格等与旅游目的地(景区景点)相关的基准关键词,利用网络搜索引擎



中的关键词推荐功能,确定本次研究的关键词为:故宫、北京故宫、故宫博物院、故宫门票、北京故宫门票、故宫门票价格、故宫地图、北京故宫地图、故宫图片、北京故宫图片、故宫博物院图片、故宫旅游、故宫开放时间共 13 个关键词。

2、数据的收集

根据《故宫博物院年鉴 2010》收集到 2010 年主要节日(春节、清明节、劳动节、端午节、中秋节、国庆节)假期期间,北京市故宫博物院所接待的游客总量。其中,十一国庆节期间的游客量在 6 个假期的游客总量中居首位,为 61.292 万人次。

百度公司推出的百度指数平台,可以提供各网络关键词在某段时间内的搜索量指数。通过百度指数平台,查询、收集到已确定关键词(共 13 个)在 2010 年五个假期期间的节日当天、提前 7 天及延后 2 天共 10 天内每天的关键词的搜索量指数。

3、百度关键词搜索量指数与实际游客量相关性分析

通过 SPSS 数据统计软件,对 13 个关键词的日搜索量指数与实际游客量进行 Pearson 相关性分析,得出每个关键词在研究时段内每天的关键词搜索量指数与实际游客量的相关性数据,其中选取相关性最大的数据列表(见表 1)。

表 1 各关键词搜索量指数与实际游客量的最大相关性数据

关键词	Pearson 相关性	显著性	时间
故宫	0.610	0.199	节日当天
北京故宫	0.596	0.212	延后 2 天
故宫博物院	0.413	0.416	提前 4 天
故宫门票	0.843	0.035*	提前 5 天
北京故宫门票	0.755	0.083	提前 2 天
故宫门票价格	0.801	0.061	提前 2 天
故宫地图	0.884	0.019*	提前 3 天
北京故宫地图	0.885	0.019*	提前 3 天
故宫图片	0.790	0.061	提前 3 天
北京故宫图片	0.479	0.336	提前 3 天
故宫博物院图片	0.759	0.080	提前 3 天
故宫旅游	0.779	0.068	延后 2 天
故宫开放时间	0.928	0.008*	提前 3 天

(*:在 0.05 水平上显著相关。)

通过分析发现,在所选取的 13 个关键词中,有 9 个关键词的搜索量指数与实际游客量的 Pearson 相关系数在 0.75 以上,相关性较大。因此,相关网络关键词搜索量指数与实际游客量之间存在较明显的相关关系,且最大相关系

数出现在节日前 2—5 天的时间范围内。在 9 个相关性较大的关键词中,有 4 个关键在 0.05 水平上显著相关,分别是故宫门票、故宫地图、北京故宫地图、故宫开放时间。其中搜索量指数与实际游客量相关性最大的关键词为“故宫开放时间”,其 Pearson 相关系数为 0.928。由此,故宫景区游客在出游前对门票价格、相关旅游路线图和开放时间的关注度较高。

四、结论

本文以北京地区的热门景区——故宫为例,基于 SPSS 的 Pearson 相关性分析方法,对所选取的 13 个关键词的搜索量指数与研究时段内的实际游客量进行了相关分析。主要得出以下结论:故宫景区在 2010 年的相关网络关键词搜索量指数与实际游客量之间存在较明显的相关关系;故宫游客对所关注的出游信息进行网络搜索查询的时间一般集中在出游前 2—5 天内;故宫景区游客在出游前最关注的出游信息主要为景区门票的价格、相关旅游路线图和景区开放时间。

基于相关性分析方法,借用信息技术提供大数据对游客量进行的相关分析,对旅游管理部门做出科学、合理的决策和安排具有一定的现实意义,为游客量的预测研究提供重要的理论依据。

【参考文献】

- [1] 百度数据研究:2014 年中国旅游百度指数报告 [EB/OL]. 中国互联网数据资讯中心,2014.
- [2] 张红贤、马耀峰:中国入境旅游市场的多元回归预测[J]. 资源开发与市场,2005(2).
- [3] 黄宏、邓婕、常彦祥、李文凤、杨东方:SPSS 在旅游人数的预测应用[J]. 山地学报,2008(51).
- [4] 李世泰、赵亚萍:基于不同时间尺度旅游客源预测模型比较研究——以山东省为例[J]. 北京第二外国语学院学报,2011(9).
- [5] 路紫、赵亚红、吴士锋等:旅游网站访问者行为的时间分布及导引分析[J]. 地理学报,2007(6).
- [6] 梁明英、王丽娜:网络旅游信息对游客行为模式影响的调查分析——以泰山游客为例[J]. 泰山学院,2008(1).
- [7] 汪正彬、赵毅:基于因子分析的自贡旅游游客满意度实证研究[J]. 乐山师范学院学报,2008(8).
- [8] 连漪、汪侠:旅游地游客量测评指标体系的研究及应用[J]. 旅游学刊,2004(5).

(责任编辑:赵小茜)