

관객개발을 위한 공연예술 소비자 빅데이터 분석 사례 고찰

김선영¹, 이의신^{2*}

¹경희대학교 공연예술학과, ²서울사이버대학교 문화예술경영학과

A Case Study on Big Data Analysis of Performing Arts Consumer for Audience Development

Sun-Young Kim¹, Eui-Shin Yi^{2*}

¹Department of Performing Arts, Kyunghee University

²Department of Culture & Arts Management, Seoul Cyber University

요약 국내 공연예술은 공급과잉과 유통영역의 부재, 뚜렷하지 않은 비즈니스 모델 등으로 인한 침체를 맞고 있다. 이러한 어려움을 타개하기 위해서는 무엇보다 객관적으로 제공되는 시장데이터를 활용해 마케팅의 효율성과 정확도를 높임으로써 관객개발과 충성도 확보가 필요한 시점이다. 본 연구는 이를 해결하기 위한 대안 중 하나가 보다 보편적이고 정확한 통계와 공연별 맞춤형 서비스 제공이 가능한 ‘빅데이터’ 분석이라는 관점에서 시작되었다. 먼저 빅데이터의 특징과 분석기술, 그리고 공연예술 소비자 분석에 대한 이론적 배경과 함께 한 신용카드사가 실시한 빅데이터 분석사례를 살펴보았다. 이를 통해 빅데이터에 의한 공연예술 소비자 연구의 의미와 한계, 그리고 그러한 한계들을 극복하기 위한 대안을 제시하고자 하였다. 사례분석 결과, 공연 구매자 대상의 카드사 데이터 자체의 불완전성, 기존 이론 검증의 한계, 낮은 활용도, 소비자 성향 및 구매 동인 분석의 한계 등이 도출되었다. 또한 이러한 문제점을 극복하기 위한 대안으로 장르와 공연명 파악이 가능하고 성향이나 구매요인 등을 추출해낼 수 있는 예매처 정보, 설문조사와의 결합과 소셜 데이터와의 매쉬업을 통해 구매동기 등의 정성적 분석을 그 대안으로 제시하였다. 이 연구는 궁극적으로는 공연예술 소비자에 대한 연구가 빅데이터 시대에 어떠한 방향으로 이루어져야하며, 어떤 변화를 모색해야 할 것인가에 대한 고민의 시작점이라고 할 수 있다. 이러한 연구결과를 바탕으로 공연예술 관객개발을 위한 보다 구체화된 정성적 분석 사례들이 많이 출현하기를 기대하며, 지금의 공연예술시장의 흐름을 정확하게 대변해 주는 빅데이터 분석과 가공을 위한 솔루션들이 계속 개발되기를 바란다.

Abstract The Korean performing arts has been facing stagnation due to oversupply, lack of effective distribution system, and insufficient business models. In order to overcome these difficulties, it is necessary to improve the efficiency and accuracy of marketing by using more objective market data, and to secure audience development and loyalty. This study considers the viewpoint that 'Big Data' could provide more general and accurate statistics and could ultimately promote tailoring services for performances. We examine the first case of Big Data analysis conducted by a credit card company as well as Big Data's characteristics, analytical techniques, and the theoretical background of performing arts consumer analysis. The purpose of this study is to identify the meaning and limitations of the analysis case on performing arts by Big Data and to overcome these limitations. As a result of the case study, incompleteness of credit card data for performance buyers, limits of verification of existing theory, low utilization, consumer propensity and limit of analysis of purchase driver were derived. In addition, as a solution to overcome these problems, it is possible to identify genre and performances, and to collect qualitative information, such as prospectors information, that can identify trends and purchase factors.combination with surveys, and purchase motives through mashups with social data. This research is ultimately the starting point of how the study of performing arts consumers should be done in the Big Data era and what changes should be sought. Based on our research results, we expect more concrete qualitative analysis cases for the development of audiences, and continue developing solutions for Big Data analysis and processing that accurately represent the performing arts market.

Keywords : Big Data, Audience Analysis, Performing Arts, Performing Arts Industry, Card Data for Performing Arts

*Corresponding Author : Eui-Shin Yi(Seoul Cyber Univ.)

Tel: +82-10-2987-4500 email: yes@iscu.ac.kr

Received November 2, 2017

Revised (1st November 24, 2017, 2nd December 7, 2017)

Accepted December 8, 2017

Published December 31, 2017

1. 서론

데이터 규모의 방대성, 데이터 처리 및 분석속도, 데이터 종류의 다양성 등을 특징으로 하는 빅데이터에 의한 분석과 인사이트(Insight) 도출에 대한 기대감이 높다. 이미 공공·행정, 의료·건강, 유통·소매, 제조업 등 다양한 분야에서 데이터 분석을 통하여 전통적인 고객관리가 아닌 다양한 미디어 속의 빅데이터를 분석·활용하여 서비스나 솔루션 개발에 적용함으로써 경쟁력을 강화하고 있다. 문화콘텐츠 부문에서도 넷플릭스의 <하우스 오브 카드>를 비롯해 <뮤직 게놈 프로젝트>, <쿠기런> 등 영상, 음악, 게임 등 분야에서 그 쓰임이 점차 확장되고 있다.

최근의 순수공연예술(본고에서의 순수공연예술은 공연법 제2조에서 정의한 공연예술 중 연예와 대중음악콘서트 등 대중예술 공연을 제외한 클래식음악, 오페라, 연극, 무용, 국악 등 분야를 의미함. 뮤지컬의 경우 대중성을 목적으로 하는 작품도 있으나 문체부의 분류기준과 순수창작뮤지컬의 비중이 높은 점, 공연예술실태조사 대상에 포함되는 점 등을 고려하여 본 정의의 범주에 포함시키하고자 함. 이하 공연예술)은 인공지능을 활용하고, 디지털 영상기술과 결합하거나 시간적·공간적·경제적 접근성을 확대하며 빠른 속도로 예술창작의 범위를 넓히고 있다. 서로 다른 표현과 소통 방식의 융합이 이루어지고 있으며 시공간적 경계를 뛰어넘어 그 영역을 확장하고 있다. 예술의 발전은 결국 기술의 활용과 지배의 과정으로, 능력과 의도, 표현매체와 표현내용의 조화된 일치로 볼 수 있다(A. Hauser, 1991)[1]. 하지만 이러한 외형적 성장에도 불구하고 현재 한국의 공연예술은 과잉공급과 유통 분야의 부재, 시장 시스템의 불확실성 등 많은 난제를 내포하고 있다. 결과적으로 공연예술의 확장은 시장의 공급과 수요의 불균형 속에 정체와 확대의 양극화 현상을 보이며 고전을 거듭하고 있는 것이다. 또한 이러한 환경 속에서 낮은 객석 점유율로 인한 수익성 악화는 예술인들의 창작활동 사기를 저하시키고 나아가 예술의 산업적 발전을 저해하는 결과를 초래하고 있다. 이렇듯 관객개발은 공연예술의 경쟁력을 뒷받침하는 주요한 수단이자 성과지표이다. 또한 공공부문이든 민간부문이든 운영에 필요한 안정적 재원이 초미의 관심사라는 것은 지극히 당연한 일이다. 하지만 지금까지의 관객개발은 개인의 감과 경험으로 관객변화를 감지하거나, 오프라인 관객 대상설문조사, 인터뷰, 대표집단 분석 등의 제한된

관객 분석으로 실질적인 관객개발과 충성도 확보에 그 한계를 보여 왔다. 이와 관련하여 이승엽(2008)은 사회적, 문화적 관행이나 인구통계학적인 특징 그리고 소득과 시간 등과 같은 것들은 공연예술분야의 관객분석에서 근본적인 한계를 가지며, 지금의 관객개발은 이런 한계 위에서 이루어지고 있다고 지적한 바 있다[2]. 예술 그 자체 뿐 아니라 산업적 요소가 공존하는 공연예술에서 객관적이고 전문적인 시장분석 모델을 발굴하고 급변하는 경제 흐름, 잦은 사회적 이슈에 대응할 수 있어야 함은 물론이다. 나아가 디지털 기술의 발달로 고객 맞춤형 서비스나 큐레이션 서비스를 요구하는 시대적 흐름에 공연예술의 유통부문도 적극적으로 대응해야 한다. 이를 위해서는 그동안 추측이나 관심도 등 비정형화된 정보를 통해서만 파악할 수 있었던 공연예술시장의 흐름을 객관적 시장 데이터라는 보다 객관적이고 명확한 언어로 그 연관성을 제시할 수 있어야 할 것이다. 이는 마케팅의 효율성과 정확도를 높임으로써 관객개발과 관객충성도 확보에 특히 중요한 부분이다.

정형데이터와 비정형데이터를 포괄하는 빅데이터는 기존의 방식에서 벗어나 보다 실증적이고 과학적인 관객 분석을 통해 공연예술의 유통 및 마케팅에 활용할 수 있는 다양한 가능성을 제시하고 있다. 예컨대, 관객들의 인구통계학적 특성별 공연 평가, 선호도, 연간 소비 행태처럼 복합적인 상관관계에 대한 조사 분석을 광범위하게 시도할 수 있다. 또한 공연 매출에서 발생한 정형(카드이용) 데이터뿐만 아니라 비정형(SNS) 데이터들을 다각적으로 분석함으로써, 그에 따른 마케팅 방안을 제시할 수 있는 것이다. 이는 정책입안자들에게도 공연예술정책을 결정하는 근거 자료로서 그 효용이 클 것이며, 공연예술 비즈니스 관계자들에게도 객관적 시장 데이터를 활용한 마케팅의 정확도를 높일 수 있도록 하여 그 효율성을 높여줄 수 있다.

하지만 이제까지 공연예술에서의 빅데이터 활용은 단지 데이터 비주얼라이제이션(visualization) 등 창·제작 분야나 일부 대형 티켓 판매 대행사의 자체데이터 분석, 몇몇 공연장의 공공데이터나 문화바우처 사업과 관련한 빅데이터 정도가 활용되고 있을 뿐이다. 또한 학문적으로도 빅데이터를 활용한 공연 관객 분석에 관한 연구는 거의 찾아보기 힘들다. 이에 본고는 최근 공연예술분야에서는 최초로 예술경영지원센터가 신한카드와 실시한 ‘빅데이터를 활용한 공연소비 트렌드 분석’ 사례를 활용

하여 소비자 분석의 가능성과 한계를 살펴보고 발전을 위한 개선점을 도출하고자 한다. 아직 선행연구가 없는 공연예술소비자 빅데이터 분석에 관한 본 연구는 공연예술 소비자에 대한 연구가 빅데이터 시대에 어떠한 방향으로 이루어져야하며, 어떤 변화를 모색해야 할 것인가에 대한 고민의 시작점이라 할 수 있다. 이 연구를 시작으로 빅데이터를 활용한 공연예술 관객개발전략을 위한 시사점을 제공하는 연구가 많이 이루어지기를 바란다.

2. 이론적 고찰

2.1 빅데이터의 특징과 분석기술[3]

2.1.1 빅데이터의 특징

가트너(Gartner)의 애널리스트 레이니(Doug Laney)가 빅데이터의 특징으로 제시한 3V는 거대한 크기(Volume), 다양한 형태(Variety), 빠른 생성·유통·분석속도(Velocity) 등을 의미하며 현재 가장 보편적인 빅데이터의 형태적 특징으로 자리 잡고 있다. 또한 여기에 데이터 처리의 복잡성(Complexity)을 추가하여 3VIC로 정의하거나 가치(Value), 유효성(Validity), 가시성(Visibility), 정확성(Veracity: 데이터의 신뢰수준) 등을 추가하여 4V로 규정하기도 한다. 데이터 규모의 방대성(Volume) 측면에서 빅데이터는 일반적으로 수십 테라바이트(TB) 혹은 수 페타바이트(PB) 이상 크기의 데이터를 가리킨다(함유근·채승병, 2013). 또한 데이터의 양은 해마다 60%씩 증가하고 있으며(한소월·이민수, 2012), 전 세계 디지털정보량은 2년마다 약 2배씩 증가하고 있기 때문에(P. Warden, 2011) 향후 정보시스템의 고도화, 모바일, 클라우드 컴퓨팅, 소셜 네트워크 서비스가 일상화되면서 더 큰 규모에서 정의될 수도 있을 것이다.

Table 1. Types of Big Data[4-15]

Type	Feature
structured	· Data stored in fixed fields · Relational databases, spreadsheets, etc.
semi-structured	· Data that is not stored in a fixed field but contains meta data or schema · XML or HTML text
unstructured	· Data that is not stored in a fixed field · Text documents and images that can analyze text, video, audio data, etc.

한편 스키 리스케(2011)는 ‘수많은 개별 요소 각각에 관한 것(고해상도)을 고빈도로 생성한 다양한 데이터’로 정의하기도 했는데 이는 데이터의 크기에 대한 연구자나 기관마다 그 기준이 달라 정확한 규정이 어려운 상황을 반영한 것으로 보인다.

데이터 종류의 다양성(Variety) 면에서 기존의 관계형 데이터베이스(RDBMS)가 처리할 수 있는 영역이 정형 데이터(structured data)에 한정되어 있었다면, 빅데이터는 정형 데이터는 물론 비정형 또는 반정형 데이터까지 처리능력을 확대한 개념이다. 정형 데이터는 스프레드시트와 같이 고정된 필드(field)에 저장된 데이터이며, 반정형 데이터는 메타데이터나 로그데이터 등을 말한다. 다만 정형 데이터는 빅데이터의 영역에 포함할 것인지 여부에 대해서는 논란의 여지가 있다. 이는 기존 데이터 산업과 빅데이터의 관계를 완전히 별개로 볼 것인지, 연장선으로 볼 것인지 등 관점의 차이에서 비롯된다. 반면 비정형 또는 비구조적 데이터(unstructured data)는 데이터마다 형식이 달라 통일된 구조로 정리하기 어려운 데이터를 지칭한다. 즉 SNS, 블로그, 뉴스 등을 통해 생성되는 데이터로서 텍스트, 비디오, 오디오 기타 다양한 형태의 데이터를 포함한다(Table 1). 비정형데이터는 향후 전체 데이터 중 약 90% 이상을 차지할 것으로 전망되고 있다(IDC, 2011)[4-15].

빅데이터는 또한 대용량의 데이터를 빠르게 처리하고 분석할 수 있다(Velocity)는 점을 특징으로 한다. 디지털 데이터의 생성 및 이동주기는 지속적으로 증가하고 있으며, 대부분은 잘 가공된 데이터가 아닌 가공되지 않은 원시 데이터이다(이규철·원희선, 2012). 따라서 이를 실시간 혹은 이에 준하여 저장·유통·수집·분석·처리할 수 있는지 여부가 관건인데 비약적으로 발전한 분산처리 등 빅데이터 기술이 이에 대한 해답을 주고 있는 것이다.

2.1.2 빅데이터의 분석기술

빅데이터의 활용은 일반 데이터와 마찬가지로 데이터 수집→저장→처리→분석→표현의 과정을 거치며 이에 따라 크게 저장·처리기술, 분석기술, 표현기술로 분류된다. 본 연구는 빅데이터의 활용 측면에 대한 논의임으로 분석기술에 한정해서 살펴보고자 한다.

빅데이터의 비즈니스적 가치를 실질적으로 만들어내는 것은 다양한 분석 기술들이다. 대표적 분석 기술로는 기계학습, 데이터 마이닝, 클러스터링, 신경망 네트워크,

회귀분석, 결정트리(decision tree), 연관분석, 자연어 처리, 시맨틱 검색, 링크 마이닝, A/B 테스트 등이 있다. 이 분석기술들은 빅데이터에만 적용되는 것이 아니고, 기존 데이터 분석에 활용되던 것들이다. 이 중 최근 소셜미디어 등 비정형데이터의 증가로 인해 텍스트마이닝, 오피니언마이닝, 소셜네트워크분석, 군집분석 등이 주목받고 있다(Table 2).

Process and technology	<ul style="list-style-type: none"> Process and technology are relatively simple Process / analysis process is formalized Cause-effect identification centering 	<ul style="list-style-type: none"> Processing complexity is very high due to various data sources, complex logic processing, and large data processing. No defined models / correlations / procedures needed to develop new and different processing methods Correlation centering Open software such as Hadoop, R, NoSQL
------------------------	---	---

Table 2. Major SNS Big Data Analysis Technology[6]

Analytical technology	Content
Text Mining	<ul style="list-style-type: none"> Extract and process useful information from non-semiotic text data based on Natural Language Processing Extract meaningful information from a large amount of text data, identify linkages with other information, and find out categories with text
Opinion Mining	<ul style="list-style-type: none"> Technology to determine affirmative, negative, and neutral preferences of stereotyped and unstructured text such as social media Use of market size forecasts, consumer reactions, and word of mouth analysis for specific services and products
Social Network Analytics	<ul style="list-style-type: none"> Technology based on Graph Theory Measures user reputation and influence based on social network connection structure and connection strength, and is used to identify users who act as hubs or hubs of word of mouth on social networks
Cluster Analysis	<ul style="list-style-type: none"> Collecting individuals with similar characteristics and ultimately using them to identify groups with similar characteristics For example, on the twitter, user groups related to interests and hobbies, such as user groups discussing photographs and cameras and users interested in games, can be classified through cluster analysis

Table 3. Differences between existing data analysis and big data analysis[4-26]

	Ordinary Data Analysis	Big Data Analysis
Amount of data	<ul style="list-style-type: none"> Terabyte level 	<ul style="list-style-type: none"> Petabyte level (at least 100 terabytes) In the case of clickstream data, customer data collection and analysis must be performed over a long period of time.
Data type	<ul style="list-style-type: none"> Formulated data centering 	<ul style="list-style-type: none"> High percentage of unstructured data such as social media data, log file, click stream data, call center log, and communication CDR log Factors that increase processing complexity

또한 기존 데이터분석과 빅데이터 분석의 차이를 요약하면 위의 Table 3와 같다. 빅데이터 분석은 처리해야 할 데이터양이 방대하고, 정형 데이터보다 SNS 등 비정형 데이터의 비중이 높으며, 처리가 복잡하여 분산처리 기술이 필요하다. 정형화되어있지 않은 모델과 절차 등으로 새롭고 다양한 처리방법 개발이 요구되는 것이다.

2.2 공연예술 소비자 분석

서문에서 밝힌 바와 같이 공연예술의 공급과잉 상태를 해결하기 위해서는 유통이 활성화되어야 하며 유통의 활성화는 소비자에 대한 정밀한 분석이 선행되어야 한다. 공연예술 소비자의 행동에 영향을 미치는 요인으로는 정치·경제·사회·기술 등 거시환경적 요인들과 역사·민족·계층·하위문화 등 문화적인 요소를 들 수 있다[7]. 그러나 개별 작품의 유통 활성화를 위해서는 이러한 거시적 요인들 보다 성격, 태도, 동기 등 심리적 요인이나 나이, 직업, 경제적 여건, 가족 상황, 생애주기에서의 위치 등 소비자의 개인적 요소에 대한 파악이 더욱 중요하다. 앤드리슨(Anderson, 1998)의 이론 ‘공연예술 수용과정’에 따르면 무관심→관심→시도→긍정적 평가→수용→확신 등 6단계로 공연예술의 관객은 진화해간다. 또한 미국 공연예술시장에서 1년간 한번이라도 공연예술을 관람한 사람은 ‘시도’ 이상의 단계로 표본분포 상 30.9%를 차지한다. 공연예술에서 개별 작품에 대한 홍보마케팅은 사실상 이들을 대상으로 이루어지며, 관심단계(22.2%) 사람들은 신규 관객 개발의 관점에서 접근한다. 1년간 공연예술 관람 경험이 전무하며, 관심조차 없는 ‘무관심’ 1단계 그룹은 47%를 차지하는데 앤드리슨은 이들을 공연장으로 이끌어내는 것은 그들의 생활양식 자체를 바꿔야하기 때문에 사실상 거의 불가능한 일이라 한다. 하지만 현대의 공연 소비자들은 공연을 소비하는 자세부터 전통적 관객과는 본질적으로 다른 변화들을 보이고 있

다. 예컨대 무관심 그룹이라고 해서 교육을 통해서만 공연소비자가 되거나 그렇지 않으면 영원히 무관심층으로 남는 것은 아니다. 1970년대 부르디외(P. Bourdieu)의 계층 간 취향의 차이를 비판한 피터슨(Peterson & Simkus, 1992)의 옴니보어(Omnivore)이론에 의하면 현대 사회에서 비단 상류계층 뿐 아니라 개인, 특히 젊은이들은 자신들의 삶에 중요한 것을 결정하는 데 매우 능숙하며 의미의 구조를 제공하는 사회계층, 종교, 국적 또는 인종에 의지하지 않는다(Fomas, Linberg and Semhede, 1995)[8]. 즉 고된 노동의 속박에서 벗어난 하류계층을 포함한 현대의 관객은 하나의 계층제(hierarchy)보다는 여러 문화적 계층과 문화들 중에서 선택한다(Kolb, 2005)[9]. 이러한 상황에서 공연예술단체나 기관이 소비자들의 기대를 충족시키기 위해서는 무엇보다 포괄적 범위의 사람들이 공연상품을 좋아하고 소비하는 이유를 이해해야 한다. 또한 신규 관객개발을 위해서는 공연상품을 소비하지 않는 원인파악이 필요하다. 현재 대부분의 공연예술단체나 기관은 객석 점유율, 관객의 정체성, 관객 수를 파악하기 위한 정량적인 인구통계조사를 실시하고 있지만, 관람동기와 관객들이 원하는 편익을 밝히려는 정성적 연구에는 관심이 적었던 것이 사실이다(Kolb, 2005)[9].

지금까지의 관객개발에 대한 연구는 크게 공연 예술 소비자의 특성에 관한 연구와 관객 세분화 및 이에 따른 관객개발 방안으로 나눌 수 있다. 영국 ACE(2011)는 영국 전역의 성인집단을 예술의 관여도를 기준으로 highly engaged, some engaged, not currently engaged 등 3개의 집단으로 나누고, 이를 다시 세부 특성에 따라 하위 13개 집단으로 분류하였다. 이어 각 하위 집단별로 인구통계학적 속성, 여가시간 활용, 미디어소비, 라이프스타일, 예술 참여요인 등을 기술함으로써 집단의 특성에 맞는 예술 참여유도 방안을 제시하였다. 결론적으로 예술 참여에 영향을 줄 수 있는 다양한 변수를 종합적으로 고려하고, 집단별 마케팅 및 홍보 전략을 구체화하였다[10].

한편 개인의 가치 및 정체성 등 다양한 변수를 활용하여 공연 관람 행태와의 상관관계를 심층적으로 분석한 연구도 있다. 공연에 관심은 있지만 실제 참여도는 낮은 사람들(interested non-attendees)을 중심으로 공연 관람의 주요 동기와 장애요소를 연구한 미국 NEA(2015)의 보고서에서는 예술의 참여는 인구통계학적 생애주기뿐 아니라 사회경제적 지위와 계급 정체성, 그리고 개인적

가치가 중요한 차이를 유발함을 보여준다[11]. 또한 남인용·백경옥(2012)의 연구에 따르면 성별, 연령, 학력, 소득에 따라 최근 관람한 공연장르 및 입장권 구매 시기, 공연 선택 시 고려요인 등이 다른 것으로 조사되었고 공연소비자의 유형(무관심관객, 관심관객, 열성관객)에 따라 선호하는 장르나 정보 획득 및 구매방법에 유의한 차이를 보였다[12].

한국 클래식 음악시장에 새롭게 등장한 관객들의 예매내역을 통해 공연소비의 패턴을 살펴보고, 설문조사로 구체적인 관람형태 및 태도에 대해 연구한 최연식(2015)은 옴니보어(omnivore) 이론의 개방성(openness)을 중심으로, 각 관객 그룹이 공통적으로 가지고 있는 대중적 특성에 주목하여 앙상블 디토 관객을 ‘클래식 대중’으로 개념화했다[13]. 한편 이용관(2012)은 서울시 소속 공연예술단체의 관객구조를 분석하여 현재 단체들의 관객기반 취약성을 살펴보고 이를 관객개발 성과가 현저히 높아지고 있는 서울시립교향악단의 관객구조와 비교해 봄으로써 관객개발 방식의 단초를 찾아보려 하였다[14]. 또한 개인의 속성이 공연 소비에 미치는 영향을 분석한 결과, 여가활용방식과 개인의 교육 수준에 따라 공연 소비가 달라지며(배영, 2013)[15], 예술에 관여도와 지식수준이 높은 집단에서 공연 관람률이 가장 높았고, 저관여 집단 및 40대 집단은 공연 관람률이 낮은 것으로 나타났다(유은영·진현정, 2014)[16]. 이처럼 지금까지 관객개발을 위한 연구는 적잖게 이루어졌으나 기존의 선행연구들이 대부분 관람동기의 분류 및 소비형태에 초점을 맞추고 있다.

공연예술 관객에 관한 선행연구를 좀 더 살펴보면 인구통계학적 분석 및 이에 따른 유형화(최운영, 2008; 박재홍, 2007), 만족도에 미치는 요인 및 이에 따른 재구매 의사 등 영향(송필석, 2008; 조선휘, 2015; 남정미, 2012; 이루지, 2006; 김보름, 2016; 강하나, 2015; 류가연, 2014; 박명숙, 2014), 작품선택에 미치는 영향 요인 및 이에 따른 유형 분류(배수연, 2005; 이화순, 2011; 이창진, 2014; 채지영, 2011; 신혜선, 2006; 현진, 2012; 권혁인, 2016; 임승희, 2014; 함현진, 2011; 주종현, 2017), 관람객 구조 비교 연구(이용관, 2012; 현진, 2012), 공연장르 간 전환행동 연구(안성아, 2012), 비관람객 유형 연구(손지영, 2016) 그리고 관객 개발에 영향을 미치는 예술교육 등 프로그램 연구(김경옥, 2015; 조인희, 2008; 홍미성, 2012; 강하나, 2015; 전태유, 2012; 이진호,

2009) 등을 들 수 있다.

이처럼 공연예술의 관객분석을 정량적 조사와 정성적 조사로 분류할 때 이제까지는 주로 설문조사 방식에 의한 정량적 분석에 의존해 왔다(신미경, 2006)[17]. 2천년대 중반 들어 ZMET기법이나 포커스그룹인터뷰(FGI) 등 정성적 조사에 의한 분석이 도입되었지만 개별 공연에 적용하는 데는 역시 제한적이라고 할 수 있다. 정량적이든 정성적이든 설문조사에 의한 관객분석은 편리하고 비용이 저렴한 반면 몇 가지 한계를 가진다. 첫째 설문지의 구성항목이 설문조사마다 거의 유사하다는 것이다. 인구통계학적·소비형태적·관람행태적·심리분석적 특성으로 구성되는 설문지는 구성순서나 질문의 심도수준이 거의 같다. 또한 설문지 작성자의 자질에 따라서 수준이 달라지며, 선택지 숫자도 제한적일 수밖에 없다. 이는 보다 다양한 견해와 의견을 반영하는데 걸림돌로 작용한다. 특히 설문지 상의 질문항목만으로 관객의 관람동인을 파악하기란 쉽지 않다. 둘째 비용·시간·인력의 제약으로 인해 조사대상이 제한적일 수밖에 없는 경우가 대부분이다. 2천년대 들어 인터넷을 활용한 설문조사를 통해 이러한 오프라인 조사의 한계를 극복하려는 시도가 활발하게 펼쳐지고 있으나 표본추출의 문제, 이메일 주소 확보 및 스팸메일로 간주되는 등의 응답률 저하 등이 문제점으로 지적되고 있다(최영훈·명승환·이강준, 2008)[18]. 셋째 같은 이유로 빈도 면에서 개별 공연에 적용할 만큼 충분히 조사가 이루어질 수 없다. 학술적 연구에 있어서는 단 일회의 설문조사만으로도 이론적 검증을 위해 충분한 경우가 많으나 공연마다 다른 결과가 나올 수밖에 없는 실제 공연현장 관객조사의 경우는 수시로 이루어져야 하지만 현실은 그렇지 못하다. 넷째 공연 전후 또는 수요인지 단계, 정보탐색 단계, 구매결정 단계로 나뉘는 공연상품 소비에 대한 결정과정에서 각 단계별로 변화되는 관객의 성향을 파악하기가 어렵다. 이 또한 일회성에 그칠 수밖에 없는 설문조사의 한계이다.

이러한 설문조사의 한계를 극복하기 위한 대안으로 매출데이터를 활용한 관객연구(전해웅, 2015)가 있으나 이 또한 예술의전당 예매시스템을 통한 구매 데이터에 국한된다는 제한점이 있다. 아울러 일반적으로 적용할 수는 있지만 개별 공연의 특수성을 반영하지 못하며 이는 설문조사도 마찬가지이다.

이상 살펴본 바와 같은 설문지 조사나 특정 공연장의 매출데이터 분석의 한계를 극복하기 위해서는 빅데이터

를 활용한 관객분석의 적극적 도입이 요구된다. 빅데이터의 정성적 분석은 다양한 연관키워드를 추출할 수 있어 작성자의 자질에 좌우되지 않는 충실한 분석이 가능하며 다양하고 심도 있는 관람동인 파악이 가능하다. 또한 표본이 확대될 수 있으며 응답의 신뢰도 문제나 응답률 저하의 문제에서도 자유롭다. 마지막으로 구매결정단계별 맞춤형 큐레이션 서비스가 가능하며 이는 관객개발에 또다른 전환점이 될 수 있을 것이다.

3. 빅데이터를 활용한 공연소비분석사례

3.1 개요

이제까지 타 분야와 달리 공연예술에서는 그 유용성에도 불구하고 빅데이터를 활용한 관객분석 사례를 찾아보기 어렵다. 이에 (재)예술경영지원센터는 공연예술의 산업적 진흥의 관점에서 (주)신한카드에 의뢰하여 2016년 카드데이터를 토대로 한 ‘공연소비 트렌드 빅데이터 분석’을 실시하였다. 이 프로젝트는 카드데이터 자료 분석을 통해 공연소비의 유형별 특성을 파악함으로써 소비자를 심층적으로 이해하는 것을 목적으로 한다. 공연예매처에서 발생한 카드데이터의 경우 대중예술공연 또는 대중음악콘서트를 제외하였고, 공연장은 공연예술실태 조사의 조사대상으로 한정하였다. 다만 순수예술을 위한 공연장이라 할지라도 대중음악공연이 일어난 일부의 경우는 카드데이터의 특성상 배제하지 못하였다. 하지만 전반적 분석에 큰 영향은 주지 못할 것으로 사료된다. 주요 내용은 ‘공연소비 트렌드 분석’, ‘공연소비 고객특성 분석’, ‘공연소비자 분류(segmentation)’, ‘상위 5% 공연소비자 이용현황 분석’, ‘국립운영 3개 공연시설 이용행태 분석’ 등인데 본고에서는 ‘국립운영 3개 공연시설 이용행태 분석’을 제외한 나머지 네 개 분석내용을 살펴보고자 한다. ‘공연소비 트렌드 분석’은 ’14년 1월~’16년 9월(3년)이며 분석기간은 ’15년 10월~’16년 9월(1년)이다. 또한 서울·경기 소재 공연시설, 예매처 등 공연 관련 가맹점(온라인/오프라인 포함)에서 신한카드를 이용한 전국거주 회원(2014년 약 104.2만 명, 2015년 약 118.9만 명, 2016년 약 82.5만 명)으로 월별, 소비자 거주지별 소비현황, 온라인/오프라인의 이용행태 비교 등을 자료 분석 범위와 대상으로 하였다. 그리고 성별·연령대별 공연소비 이용행태, 성별·연령대별·세부거주지별

(서울·경기) 공연소비 이용행태, 성별·연령대별·직군별 공연소비 이용행태, 성별·연령대별·소득구간대별 이용행태 등을 분석하였다.

‘공연소비자 분류(segmentation)’는 서울·경기소재 공연 관련 가맹점(온라인/오프라인 포함)에서 신한카드를 이용한 서울·경기 거주자 중 공연 외 4개 업종을 추가 이용한 회원(약 3.8만 명)을 대상으로 하였다. 또한 ‘상위 5% 공연소비자 이용현황 분석’은 서울·경기 소재 공연 관련 가맹점 (온라인/오프라인 포함)에서 신한카드를 이용한 전국 거주 회원 약 5천 명을 대상으로 하였다. 분석 프로세스는 아래 Figure 1과 같다.

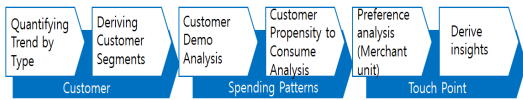


Fig. 1. Analysis process[19-7]

3.2 분석 결과 요약

3.2.1 공연소비 트렌드 분석

최근 3개년 간(2014~2016년) 카드매출 데이터를 통해 공연소비 트렌드를 분석한 결과, 매년 공연소비 규모는 4% 가량 증가하고 있으며 이용건수는 3분기, 이용금액은 4분기에 가장 높게 나타났다(Figure 2).

공연소비 금액의 전국 내 비중은 서울·경기 거주자가 3년 동안 지속 감소('14년, 66.8% → '16년, 60.8%)한 반면 5대 광역시 거주자는 지속 증가('14년, 16.2% → '16년, 18.9%)하는 것으로 분석되었다. 온라인/오프라인을 비교해 보면 상반기 온라인 공연소비 결제금액은('14년) 2,050억→('15년) 2,464억→('16년) 2,455억 원으로 2015년에 크게 성장했고 이용 회원수 비중 또한('14년) 76.1%→('15년) 79.2%→('16년) 82.1%로 증가하는 경향을 보였으며 2016년 기준 온라인 결제금액의 비중은 88%에 달했다(Figure 2).

온라인 이용횟수가 가장 많은 집단은 40대 초반의 남성(83.6%)과 30대 초반의 여성(88.1%)이며, 오프라인 이용이 가장 많은 집단은 60대 이상의 남성(24.7%)과 여성(26.3%)으로 나타났다(2016년 상반기 기준).

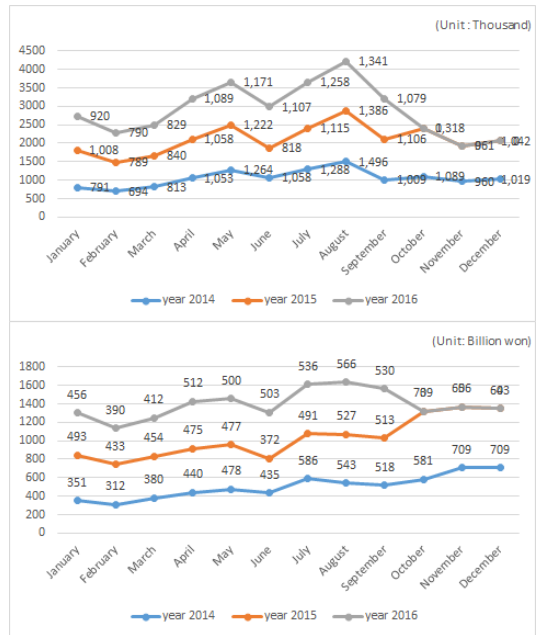


Fig. 2. In recent 3 years, the number of payments (top) and amount of money (down)[19-5]



Fig. 3. Trends in On / Offline Payments (top) and Frequency (down) in Recent 3 Years[19-20]

3.2.2 공연소비 고객특성 분석

최근 1년간 연령별 공연소비 이용금액은 30대 초반(1,016억)→40대 초반(771억)→30대 중반(768억)→20대 후반(724억) 순으로 나타났으며 60대 이상 연령대(191억 원)에서 가장 낮았다. 남성의 경우 30대 초반(393억)

→40대 초반(339억)→ 30대 중반(321억) 순으로 전체와 동일한 순서로 나타났으나, 여성의 경우 30대 초반(623억)→20대 후반(492억)→30대 중반(447억) 순으로 비교적 젊은 연령대의 이용금액이 많았다.

특히 연령대별 현황에서 30대 초반~40대 초반 연령대가 전체 이용금액 비중의 50.8%로 공연소비를 주도하고 있으며, 50대의 공연소비도 9.4%로 적지 않은 모습을 보였다. 또한 Table 4와 같이 20대 후반~30대 중반 구간에서 일인당 평균 이용금액이 각 23만2천원, 23만7천원, 20만원으로 다른 연령대에 비해 높게 나타났다.

Table 4. Card Usage for Performance Consumption by Age Group[19-22]

(Unit: thousand cases, thousand persons, 100 million won) (Unit: case, thousand won)

Age group	Number of Payments	Members using	Amount used	%	Average Number of payments	Amount used by case	Amount used by person
Early 20s	704	211	245	3.9	3.3	35	116
Mid 20s	1,000	277	465	7.4	3.6	47	168
Late 20s	1,223	312	724	11.5	3.9	59	232
Early 30s	1,686	429	1,016	16.1	3.9	60	237
Mid 30s	1,457	385	768	12.2	3.8	53	200
Late 30s	1,416	377	642	10.2	3.8	45	170
Early 40s	1,762	459	771	12.2	3.8	44	168
Mid 40s	1,108	299	510	8.1	3.7	46	171
Late 40s	749	210	378	6.0	3.6	50	180
50s	1,144	333	590	9.4	3.4	52	177
More than 60s	390	114	191	3.0	3.4	49	168
Sum	12,640	3,406	6,298	100	3.7	50	185

거주지별로 보면 전국 17개 시도 중 서울·경기지역 거주자의 이용금액 비중이 61.0%로 가장 높았는데 이는 서울·경기 소재 공연시설을 주요 대상으로 했기 때문인 것으로 풀이된다. 서울은 강남구(8.3%), 송파구(7.6%), 서초구(7.0%) 순이며, 경기도는 성남시(13.4%), 수원시(10.2%), 고양시(10.0%) 순으로 높은 이용금액 비중을 보였다.

직군별로는 일반기업과 연금·금융소득 직군의 비중이 전체의 56.3%로 대다수를 차지하며 교사(6.5%), 의사(4.8%), 공무원(4.6%)이 그 뒤를 이었다. 직군별 이용금액 비중에서는 남성의 경우 군인과 경찰의 비중이, 여성은 학생과 주부의 비중이 각 상대성별에 비해 높게 나타

났다.

소득구간대별 이용현황을 보면 월 평균 소득 700만 원 이상 구간의 인당 이용금액이 26만원으로 가장 높았으나, 500만 원 이하 구간에서는 소득이 증가할수록 인당 이용금액이 낮아지는(19만6천원 → 17만7천원) 현상을 보였는데 이는 이른바 린더효과(Linder Effect)를 입증해주는 것이라 할 수 있다.

3.2.3 공연예술 소비자 분류(Segmentation)

공연예술 소비자 분류에는 R.F.M 분석과 결정트리(Decision Tree) 등의 분석기법을 사용하였다. R.F.M(Recency, Frequency, Monetary)분석은 최근 1년간(Recency) 공연소비횟수(Frequency) 및 이용금액(Monetary)을 기준으로 한 세그먼트(segment) 분류 및 소비행태 분석이며, 결정트리는 공연소비와 관련된 데이터(성별, 지역별, 연령별, 이용횟수, 이용금액, 공연소비 이외의 업종별 이용유무, 이용비중 등)를 변수로 사용하여 공연소비자를 최고액~최저액 소비층으로 분류하는 데이터마이닝 방법론이다.

분석내용으로는 segment별 공연이용행태 분석, 공연 소비자 유형별 연관소비 분석 등인데 이때 연관소비 업종 분류는 한국문화관광연구원의 ‘국민여가부문 신용카드 지출액 분석’ 업종분류 기준에 따른다.

분석결과 Table 5와 같이 공연소비자는 총 10개의 Segment로 분류되었다. 공연소비에 가장 많은 지출을 하는 <seg. 1>은 평균소득 150~270만원, 20대 후반~30대 초반, 학생/교사, 서울 강서구/광진구/관악구 거주 여성으로 구성되었다. 대체로 여성이 남성에 비해 공연 소비 금액이 높았으며, 소득수준에 관계없이 개인 여가 시간이 많은 직군이 높은 공연소비 성향을 보였다.

최고액 소비 고객군과 최저액 소비 고객군의 일인당 이용금액은 최고 48만 8천원, 최저 26만 6천원이고, 건당이용금액은 최고 8만원, 최저 4만1천원으로 약 2배의 차이를 보였다. Segment별 연관소비와 주요소비처를 분석한 결과(Table 6), 문화생활 연관소비(영화, 미술관 관람 등)가 많을수록 공연소비 성향이 높으며, 스포츠와 같이 활동적이고 아웃도어 지향적인 취미활동(스포츠센터, 골프 등)을 가질수록 공연소비 성향이 낮은 것으로 나타났다.

Table 5. Performance Consumer Segment Classification[19-55]

Seg.	ratio (%)	gender	Average income (Ten thousand won)	Average age range	Main occupation	Primary residence
1	15.2	F	150 ~270	Late 20s ~Early 30s	Student, teacher, academy lecturer	Gangseo-gu, Gwangjin-gu, Gwanak-gu
2	11.8	F	250 ~350	Early 40's ~ 50s	Pension / Finance Income, Housewife	Gangnam-gu, Seocho-gu, Yangcheon-gu
3	6.8	F	250 ~350	Mid-thirties	Housewife, Financial institution, doctor	Seocho-gu, Gangnam-gu, Songpa-gu
4	5.3	F	210 ~290	Early 30's	doctor, academy lecturer	Seongbuk-gu, Seongdong-gu, Yeongdeungpo-gu
5	14.5	M	200 ~380	20's ~Late 40s	General business, military / police, student	Gwanak-gu, Paju City, Suwon
6	5.6	F	230 ~325	Mid-thirties ~Early 40s	Housewife, academy lecturer, teacher	Yongin, Suwon, Seongnam
7	7.2	M	240 ~360	Mid-30s	General business, military / police, student	Hwaseong, Suwon, Anyang
8	15.6	M	325 ~550	Early 40's ~ 60s	Public servants, other manufacturing, wholesale and retail	Gangnam-gu, Goyang, Yongin
9	8.4	M	290 ~480	Late 30s ~Early 40s	General corporate, professional, media	Songpa-gu, Mapo-gu, Seocho-gu
10	9.6	M	290 ~440	Mid-thirties ~Early 40s	General business, military / police, professional	Yongin, Suwon, Hwaseong

Table 6. Performance Consumption and Associated Consumption Behavior by Segment[19-61]

Seg.	Number of payments	Person Use amount	Number Use amount	First priority	Second priority	Third priority	Main consumption place
1	6.1	487,739	79,913	Beauty (+28.8%)	Cultural life (+9.7%)	Hobbies (+6.1%)	Public transportation, coffee /

				%)			dessert, CVS
2	5.6	392,337	70,549	Educational (+32.5%)	Hobbies (+26.5%)	Beauty (+11.5%)	Supermarket, Monthly payment (administrative expenses such as electricity), Dept. store
3	7.0	389,659	55,716	Educational (+55.9%)	Hobbies (+2.0%)	Travel (-2.3%)	Online shopping, public transportation, coffee / dessert
4	6.4	368,266	57,187	Travel (+13.0%)	Beauty (+11.9%)	Cultural life (-0.9%)	Online shopping, Medical, department store, public transportation
5	5.4	359,545	66,263	Hobbies (+8.7%)	Cultural life (+8.2%)	Travel (+3.1%)	Convenience store, Gasoline, Eat out, Coffee / Dessert
6	6.0	338,043	56,334	Educational (+52.8%)	Hobbies (+1.1%)	Sports (+0.5%)	Mart, Online shopping, Life-friendly consumption
7	5.8	322,501	55,453	Travel (+21.2%)	Cultural life (+10.1%)	Sports (+6.3%)	Convenience store, High pass Maintenance, Eat out
8	5.6	304,148	54,046	Educational (+31.2%)	Sports (+25.8%)	Hobbies (+3.5%)	Supermarket, Fueling, Education, Dining out, Golf
9	6.8	294,806	43,474	Sports (+10.9%)	Travel (+9.3%)	Educational (+2.0%)	Convenience store, Medical, Taxi, Public transportation
10	6.5	265,647	40,973	Sports (+28.7%)	Travel (+3.9%)	Educational (-0.1%)	Mart, Medical, Gasoline, Month management fee, Sports, Golf

3.2.4 상위 5% 공연소비자 이용현황

상위 5% 소비자 중 이용금액 비중이 가장 높은 고객군은 30대 초반의 여성으로 전체 이용금액의 8.1%를 차지하며, 비중이 가장 적은 고객군은 20대 초반의 남성으로 전체의 0.3%를 차지했다. Figure 4에서처럼 일인당 이용금액은 남성이 288만원으로 여성의 248만원보다 약 40만원 많지만 상위 5% 소비자 내 여성 회원수가 3,867명(남성 5,220/여성 9,087) 더 많기 때문에 전체 이용금

액은 여성이 225억 원으로 남성의 150억 원보다 약 75억 원 더 많았다.

상위 5% 소비자의 거주 지역별 비중은 서울→경기→인천→대구→경남 순으로 전체 소비자의 비중 순인 서울→경기→인천→경남→대구와 크게 다르지 않았다. 또한 상위 5% 소비자의 소득별 인당 이용금액은 700만원 초과 소득군을 제외하고 소득과 비례하는 경향을 보였다. 이는 전체 소비자의 소득별 인당 이용금액현황에서 200만 원 이하 소득구간이 다른 소득 구간대와 비교해 큰 금액을 소비했던 것과 차이가 있다.

상위 5%인 공연소비자의 경우 대조군 대비 문화 활동 소비 비중이 높고, 공연관람의 대체재 소비에는 상대적으로 소극적인 모습을 보였다.

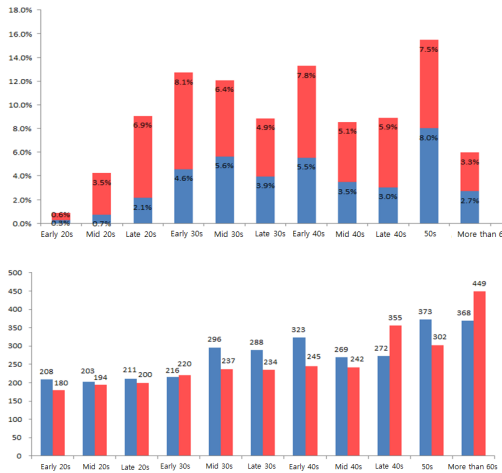


Fig. 4. Top 5% Performance Consumer's share of usage amount (up) and per capita usage amount (bottom) by gender, age[19-59]

4. 사례 분석

4.1 의미

사례로 제시한 ‘공연소비 트렌드 빅데이터 분석’은 공연예술 분야에서 카드결제 건수에 대한 빅데이터 분석을 시도한 것으로, 그간 표본추출을 통한 조사에 의존했던 한계를 극복하고자 하는 시작점이라는 것에 큰 의미가 있다. 특히 상위 5% 공연소비자에 국한되기는 했지만 연관소비분야 및 주요소비처의 분석을 통해 공연관객의 소비행태를 도출해보고자 한 것은 카드데이터의 특징을

살린 시도였다. 또한 부수적으로 파생된 결과지만 다양한 공연예술 소비자와 관련된 시간의 이동을 광범위한 데이터를 통해 검증했다는 측면도 고무할 일이다.

먼저 공연시장의 주요 소비자는 일반 소비자보다 다양한 문화 활동 부문에서 적극적 소비를 한다는 점을 보여줌으로써 피터슨(Peterson & Simkus, 1992)의 옴니보어(Omnivore)이론을 일부 실증하였다. 또한 1970년대에 제기되었던 린더효과(Linder Effect)를 실증적으로 입증했다는 점에서도 큰 의미가 있다. 즉 소득이 증가한다고 해서 공연 소비액이 함께 늘지 않고 일정 구간(300만 원대)에서 가장 많은 소비가 발생함으로써 공연소비에 있어서는 소득보다 시간적 여유가 중요한 변수임을 밝혔다. 아울러 매니아 층에 해당하는 상위 5% 소비자에게는 린더효과가 적용되지 않음을 밝혀낸 점도 그 성과라 할 수 있다.

Table 7. Contrasted group and top 5% Consumer consumption ratio comparison[19-66]

	Contrasted group	top 5% Consumer	GAP
Travel	35.7%	38.1%	6.7%
Beauty	21.7%	20.3%	-6.1%
Cultural activity	14.3%	15.9%	10.7%
Sports	13.1%	12.1%	-7.5%
Hobbies	10.6%	9.5%	-9.8%
Education	4.6%	4.0%	-12.5%
Sum	100.0%	100.0%	

사례는 또한 앤드리슨의 공연예술 수용과정 단계의 정의가 국내의 현재 상황과 일부 다를 수도 있다는 점을 밝혀냈다. 사례에서 분류한 상위 5% 소비자는 앤드리슨의 공연예술 수용과정 중 6단계인 ‘확신단계’에 해당한다고 할 수 있다. ‘확신단계’는 1년간 6개의 공연예술부문(재즈, 고전음악, 오페라, 뮤지컬, 발레, 연극 등) 중 하나 혹은 그 이상에 두 번 이상 참석할 사실이 있으며 더 많이 참석하는 데 관심이 있는 그룹이다. 앤드리슨은 확신단계에 있는 공연소비자(포본 분포의 8.7%)는 ‘여러 예술단체에 대한 정기회원이거나 여러 예술단체의 주소록 혹은 회원목록 등록자’, ‘역사유적지 방문자’, ‘운동강좌나 헬스클럽 참여자’, ‘여성’, ‘대도시 거주자’ 등으로 구성된다고 하였다. 그러나 사례의 상위 5% 연관 소비를 보면 ‘역사유적지 방문자’의 경우 ‘여행(온천, 숙박, 면세점, 수족관, 민속촌, 항공사, 고속버스 철도 여객선,

렌터카 등) 업종과 유관된 것으로 간주할 때 그 연관성을 가지나, ‘스포츠(스포츠센터, 스키장, 볼링장, 수영장, 골프, 당구 등)’나 ‘교육(일반전문학원, 어학원, 보습학원 등)’ 업종에서 대조군 대비 비중이 낮다. 이러한 측면에서 앤드리슨의 ‘운동강좌나 헬스클럽 참여자’는 우리 현실과 거리가 있다고 할 수 있을 것이다(Table 7).

이러한 이론의 검증 외에도 계절에 따른 공연소비행태를 추측할 수 있는 유의미한 데이터가 제시되었다. 카드 이용금액이 성탄절과 연말행사가 몰려있는 4분기에 가장 많은 것과는 대조적으로 이용건수 면에서는 공연의 비수기로 알려진 8월에 가장 많았다. 이는 방학을 맞은 자녀들과 함께 저렴하게 관람할 수 있는 공연 소비가 많아지기 때문으로 분석된다. 4분기의 경우 이용금액이 많은 반면 이용건수가 상대적으로 적은 것은 연말을 맞이해 고가의 대형 작품들을 관람하는 것으로 추측해 볼 수 있다.

4.2 사례분석의 한계

제시한 사례는 매출규모의 추이분석이나 소비자의 그룹화에는 소정의 목적을 달성했으나 소비자의 특성이나 소비성향을 밝혀내는 데는 한계를 드러냈다. 공연의 투자 유치나 마케팅 전략 수립, 공연 제작을 위한 자료로 활용하는데 있어서 사례의 제한점을 살펴보면 다음과 같다.

첫째, 특정 카드사 고객만을 대상으로 했다는 점, 전국이 아닌 서울·경기 소재 공연시설, 예매처 등에 국한되었다는 점에서 전수조사로서의 완벽성은 결여되어 있다고 할 수 있다. 카드사별로 고객의 성향이 다를 수 있으며, 다르지 않다면 이를 먼저 입증하는 작업이 선행되어야 한다. 특히 구매자와 실제 관람자가 다를 수 있다는 측면에서 신뢰도의 문제가 발생한다.

둘째, 장르별·작품별 소비데이터에 대한 분석이 이루어지지 않아 광범위한 데이터를 분석했음에도 불구하고 활용도가 떨어진다. 그중 하나가 코호트효과(Cohort Effect)를 입증할 수 없었다는 점이다. 코호트는 5년(예: 1960~65년) 혹은 10년(예: 1970~80년) 사이에 태어난 사람들을 말할 때 사용되며 사회변동 연구에서 특히 중요한 개념이다. 연령증가와 더불어 자연스럽게 나타나는 종단적(縱斷的) 효과와 달리(고영복, 2000) 각각의 ‘코호트’는 새로운 사회적 경험을 서로 다른 역사적 상황에서 각자의 방식으로 하기 때문에 문화적 가치나 신념 등을 재해석하여 사회적 변동에 기여한다(전해웅, 2015)[20-64].

따라서 코호트의 분석은 공연예술소비자의 세대계층 간 성향을 파악할 수 있는 틀이 되지만 본고의 사례를 통해서 할 수 없다는 점은 아쉬운 부분이다.

또한 피터슨의 옴니보어 이론을 실증함에 있어서도 불완전성을 보였는데 이는 연관소비영역의 ‘문화활동’ 분류의 내용에서 기인한다. 사례의 연관소비와 관련한 분류는 한국문화관광연구원(2016)의 ‘국민여가부문 신용카드 지출액 분석 상세업종 구분’을 따랐는데 여기에서 ‘문화 활동’은 ‘음반, 악기, 피아노, 사진기, 극장’ 등을 내용으로 한다. 하지만 피터슨의 옴니보어(omnivore, 포괄적 향유)개념에 따르면 현대에는 하이브로우 스nob(highbrow snob: 고상한 배척자)에서 고급문화와 대중문화를 동시에 향유하는 특성으로 변화되고 있다(류승완 외, 2016)[21]. 따라서 사례에서 기준으로 삼은 ‘문화활동’의 범주를 살펴볼 때 다양한 문화활동을 즐긴다는 측면에서 ‘잡식성’은 맞지만 고급문화와 대중문화를 동시에 향유하는 피터슨의 ‘포괄적 향유자’로 범주화하기는 어렵다. 한국 사회는 사회계급보다는 연령과 성별로 대표되는 인구 통계적 요소가 훨씬 중요하다는 다수의 연구결과(전범수·이상길, 2004; 조수현, 2005; 양아름, 2012; 남인용·백경옥, 2012; 이용관·장구보, 2012)를 실증하지 못한 것이다.

셋째, 무엇보다 가장 큰 한계는 공연소비자의 성향분석 내지는 소비자의 행동 요인을 도출해 내기가 어렵다는 점이다. 즉 소비자의 관람 이유와 경험에 대한 만족을 조사하는 동기 조사(motivation research, Oppenheim, 1992)가 거의 이루어지지 않은 것이다. 이는 사례의 분석 대상이 카드매출이라는 정형 데이터에 국한되었기 때문이다. 구매 데이터 분석은 기억력의 한계와 응답자의 심리적 변수, 응답의 정직성 여부 등으로 인해 그 내용에 완전한 신뢰성을 기대할 수 없는 단점을 지닌 설문조사와 달리 실제 행동으로 나타난 소비자의 행태를 정확하게 측정할 수 있다는 측면에서 장점이 있다. 하지만 구매 데이터로 소비자의 소비행동의 방식과 결과는 정확하게 보여주지만 그 행동의 원인이나 동기는 설명해주지 않는다는 한계를 가진다. 시장세분화의 기준인 성별, 미학적 관심, 편익, 시기, 충실도, 준비도(김주호·용호성, 2002) [22] 중에서 미학적 관심, 편익, 충실도, 준비도 등도 정형 데이터로는 파악하기 어려운 것들이다. 사례에서 시도했던 연관소비 또한 구체적인 구매품목이 확인되지 않으면 소비자의 성향을 파악하는 데 실질적인 도움이 되

지 않는다. 이처럼 사례는 기존 이론이나 설문조사 등을 통해 만들어진 가설들을 입증, 수정, 보완하는 데 활용될 수는 있으나 고객이 가지고 있는 내면의 특성이나 추구하는 가치와 같은 복잡한 인지구조를 파악하는 데에는 그 한계가 있으며 특히 개별 공연의 마케팅에 활용되기가 어렵다.

4.3 한계 극복을 위한 제언

맥킨지(McKinsey, 2012)는 빅데이터의 사회경제적 가치로 산업의 투명성 증대, 소비자 니즈 발견·트렌드 예측 성과 향상, 소비자 맞춤형 비즈니스 및 고객 세분화, 자동 알고리즘을 통한 의사결정 지원과 대행, 비즈니스 모델·상품·서비스 혁신 등 5가지를 제시하였다[23]. 결국 빅데이터는 데이터마이닝을 통해 새로운 인사이트를 찾는 ‘지식발견 단계’, 그리고 최종적으로 ‘의사결정단계’에 이를 수 있을 때 그 진정한 가치가 발현된다(장영재, 2012)[24]. 하지만 사례와 같이 카드매출데이터 분석만으로는 새로운 지식발견의 단계와 의사결정단계를 위한 인사이트(insight)를 도출하기가 어렵다. 더욱이 카드매출 데이터는 크기(volume)의 측면에서는 빅데이터라 할 수 있을지라도 정형·반정형·비정형 데이터를 포괄하는 형태적 의미에서의 온전한 빅데이터라고 하기에는 논쟁의 여지가 많다. 특히 공연 소비자의 특성이나 소비성향을 도출해내기 위해서는 카드데이터 등 정형적 데이터에 대한 분석 외에 소비행동요인에 대한 추가적 분석이 이루어져야 할 것이다. 이와 같은 여러 한계들을 극복하기 위해서 본고는 아래와 같이 제안하고자 한다.

첫째, 예매처의 정보를 결합하는 것이다. 카드 데이터에는 품목 정보가 없기 때문에 어떤 장르의 무슨 공연을 관람했는지 알 수가 없다. 이러한 단점을 보완하기 위해서는 장르와 공연명 파악이 가능한 예매처 정보와의 매쉬업(mash-up)이 요구된다. 하지만 여기에는 개인정보에 대한 문제, 비용과 시간의 문제를 해결해야 한다는 문제가 있으며 설령 가능하더라도 구매동기 및 반응을 도출해내기에는 여전히 한계가 있다.

둘째, 설문조사와의 결합이다. 카드데이터에 따라 소비자를 유형별로 그룹화하고 설문조사분석 시 이를 적용하면 각 그룹별 성향이나 구매요인 등을 추출해낼 수 있다. 하지만 이 경우 신뢰성과 표본수의 제한 등 설문조사 자체의 한계는 여전히 존재하고 인위적으로 규정된 집단의 성향을 파악하는 데 그칠 수 있다.

셋째, 인터넷 및 소셜데이터와의 결합이다. 공연예술은 그동안 경험재로 정의되어 왔으나, 인터넷과 모바일, 사회관계망서비스(SNS) 등을 통해 정보의 공유가 가능해지면서 탐색재의 속성이 증가하고 있다. 스마트폰의 보급 확대와 함께 SNS의 이용률이 급증함에 따라 관객의 구매의사결정 단계에서 정보탐색 비중이 점차 확대되고 있는 것이다(류승완 외, 2016)[21]. 실제로 공연 관객의 SNS이용은 급속히 확대되고 있다. 2016년 5월에 실시한 ‘SNS를 통한 공연정보 이용실태’에 따르면 77%가 페이스북, 76.7%가 트위터, 70.8%가 인스타그램을 사용하고 있다고 응답하였다(더 뮤지컬, 2016)[25]. 즉 SNS 데이터의 분석만으로도 공연 소비자의 특성을 파악하는 데 지금까지와는 다른 정밀한 인사이트를 얻을 수 있는 것이다. 우선 특정 공연의 관람요인에 대한 정성적 분석이 가능하다. 공연 시작 이전과 이후 시점별로 주로 언급된 키워드와 연관어 등 소셜 반응을 분석하면 공연 중 인기 있는 장면을 비롯한 공연관람요인을 추출해 낼 수 있다. 아울러 SNS별로 버즈(buzz)량의 추이를 비교 분석함으로써 인지 경로의 분석도 가능하다. 또한 마케팅 전후 버즈량의 추이와 긍정·부정어를 분석하면 마케팅으로 인한 인지도 확산 추이 및 마케팅의 전파력을 가늠할 수 있을 것이다.

나아가 주요 언급 키워드와 연관어를 분석하면 육하원칙 기반의 정성적 주요 속성을 빈도 및 비중별로 도출해 낼 수 있다. 즉 소셜데이터와 카드데이터를 매쉬업하면 보다 정밀하게 각 그룹별, 구매공연별, 구매시간별 정성적 구매요인에 대한 분석이 가능할 것이다. 예를 들어 ‘영포티’, ‘혼공족’ 등 최신 트렌드를 카드데이터로 확인함과 동시에 소셜데이터를 통해 그들의 정서와 취향을 추출해 낼 수 있다.

5. 결론 및 향후 과제

침체를 맞고 있다고 할 수 있는 한국의 공연예술을 위해서는 객관적으로 제공되는 시장데이터를 활용할 마케팅의 효율성과 정확도를 높이고, 관객개발과 충성도 확보가 시급하다. 또한 그동안의 비정형화된 정보 또는 제한된 설문조사를 통해서만 파악할 수 있었던 공연예술의 시장 흐름을 데이터라는 보다 객관적이고 명확한 언어로 그 연관성을 제시할 수 있어야 한다. 이를 위한 가

장 큰 대안은 ‘빅데이터’라는 생각으로 시작된 본 연구는 공연예술 소비자에 대한 연구가 빅데이터 시대에 어떠한 방향으로 이루어져야하며 또한 어떤 변화를 모색해야 할 것인가에 대해 고찰하고자 하였다. 사례로 제시한 ‘공연소비 트렌드 빅데이터 분석’ 결과, 공연 구매자 대상의 카드사 데이터 자체의 불완전성, 기존 이론 검증의 한계 등의 문제점 등이 도출되었다. 물론 본 연구에서 제시한 사례는 특정 1개 카드사의 매출데이터에 한정되지만 해당카드사와 제휴함으로써 할인이나 서비스 등의 혜택을 받는 공연만이 아닌 모든 공연을 대상으로 한다. 다만 카드사별로 이용자의 취향이나 특성이 미세하게 다를 수 있기 때문에 일반화에 문제가 완벽하게 없다고 단언하기는 어렵다. 이 부분에 관한 심층 연구는 추후의 과제로 남겨두기로 한다. 흔히 빅데이터 분석하면 스케일이 큰 거시적 측면의 인사이트를 기대하는 경향이 많다. 하지만 빅데이터의 가치는 단순히 많은 데이터량을 의미하는 것이 아니라 이를 통해 새로운 가치를 창출해내는 데에 있다. 또한 빅데이터 분석은 아직 초기단계로서 개인정보 보호 문제 등 선결되어야 할 과제들로 당장 거창한 인사이트를 도출해내기가 쉽지 않다. 무엇보다 소셜 가입자의 성별, 연령 등 개인정보는 API가 공개되지 않아, 빅데이터 업체가 크롤링(crawling) 할 수 없다. 2016년 6월에 개인정보 비식별 조치를 위한 가이드라인이 제시되었으나, 법적 효력이 없기 때문에 여전히 빅데이터 활용은 제한적이다. 그러나 한편으로는 주어진 여건 내에서 작은 흐름들을 예측하는 것 또한 의미 있는 작업일 수 있다. 예를 들어 특정 공연의 티켓판매 전후 반응의 차이를 소셜데이터나 인터넷 데이터로 분석해 내는 것만으로도 큰 성과라고 할 수 있다. 비용의 문제와 함께 기획자 등 관계자들이 관객이 원하는 것을 이미 알고 있다고 여기거나, 자신의 관람동기가 관객의 관람동기와 동일하다고 했던(Kolb, 2005)[9-43] 이제까지의 관행을 탈피하는 계기가 될 것이기 때문이다. 무엇보다 빅데이터 분석에 앞서 선행되어야 할 것은 일반적 마케팅 조사와 같이 알고 싶은 것은 무엇인지와 결과를 어떻게 활용할 것인가(Gordon and Langmaid, 1988)[26] 등 기초적 질문에 대한 해답이다. 특히 공연예술 소비자 연구에 있어서 이러한 가치를 제대로 활용하기 위해서는 기존의 정량적 데이터 뿐 아니라 소셜데이터 등 비정형데이터와의 매쉬업이 필요하다. 이를 위한 예매처 데이터, 설문조사 또는 소셜 데이터와의 매쉬업을 통해 구매동기 등의 정성적

분석을 그 대안으로 제시하였다. 이러한 연구결과를 바탕으로 공연예술 관객개발을 위한 보다 구체화된 정성적 분석 사례들이 많이 출현하기를 기대한다. 또한 향후 공연예술시장의 흐름을 정확하게 대변해 주는 빅데이터 분석과 가공을 위한 솔루션들이 계속 개발되기를 바란다.

References

- [1] A. Hauser, Trans. N. C. Baek, M. W. Yum, *Sozialgeschichte der Kunst und Literatur*, p. 119, Changbi Publishers, 1991.
- [2] S. Y. Lee, *How is the Performing Arts Market Moving?*, p. 3, Korea Arts Management Service, 2008.
- [3] S. Y. Kim, B. J. Ahn, "Big Data Public Service Model for Content Industry", *The Korea Journal of Sports Science*, vol. 26, no. 1, pp. 599-615, 2017.
- [4] S. Y. Kim, "A Study on Public Big Data Service Platform Model for Content Industry", *Chungang Univ.*, p.15, 25, 26, 2015.
- [5] S. H. Lee, D. W. Lee, "Current Status of Big Data Utilization", *Journal of Digital Convergence*, vol. 11, no. 2, 2013.
- [6] S. W. Cho, *Analysis technology in the digital age*, p.4, KT Institute of Technology, 2011.
- [7] J. H. Kim, H. S. Yong, *Arts Management*, Gimm-Young Publishers, pp. 112-113, 2002.
- [8] H. J. Jang, H. M. Jang, J. B. Kim, "Analysis of Culture Voucher from Omnivore Perspectives", *The Journal of Humanities and Social science*, vol. 7, no. 5, p. 945, 2016.
DOI: <http://dx.doi.org/10.22143/HSS21.7.5.49>
- [9] Bonita M. Kolb, *Marketing Cultural Organizations*, p. 42, Gimm-Young Publishers, 2005.
- [10] Arts Council England, *Arts Audiences: Insight*, pp. 79-83 ACE, 2011.
- [11] National Endowment for the Arts, *When going gets tough: Barriers and Motivation Affecting Arts Attendance*, pp. 134-140, NEA, 2015.
- [12] I. Y. Nam, K. O. Baek, "Watching Behavior and Media Use of Performing Arts Consumers", *KCTI*, vol. 21, pp. 5-29, 2012.
- [13] Y. S. Choi, "Suggestions for Audience Development on Classical Music Mass Audience : Focusing on the Characteristics of Omnivore in Ensemble DITTO Audience", *Journal of cultural policy*, vol. 29, no. 1, pp. 59-77, 2015.
DOI: <http://dx.doi.org/10.16937/jcp.29.1.201501.50>
- [14] Y. K. Lee, K. B. Jang, "The Audience Development Method through Comparison of the Audience Structure", *KCTI*, vol. 21, no. 21, pp. 61-89, 2012.
- [15] Y. Bae, "A Study on the Factor Effects to the Consumption of Cultural Content in Leisure Activities", *Institute of Social Research, Korea University*, vol. 14,

no. 1, pp. 73-100, 2013.

- [16] Y. E. Yu, H. J. Jin, "Analysis for Consumer Segmentation Using the Decision Tree Model in Performing Arts Market", *Journal of Consumer Studies*, vol. 25, no. 6, pp. 65-90, 2014.
- [17] M. K. Shin, "Study on Audience Research of the Field of Performing Arts in Korea", *Chugye Univ.*, p. 6, 2006.
- [18] Y. H. Choi, S. H. Myung, K. C. Lee, "The Possibility and Issues of Internet Survey II: 2001 and 2006 Survey Results", *Korean Public Administration Review*, vol.42 no.3, pp. 423-440, 2008.5.
- [19] KAMS, Shinhan Card Co., Utilizing Big Data Performance Consumption Trend Analysis Report, KAMS, Shinhan Card Co., 2017.
- [20] H. W. John, "Study on Patterns of Performing Arts Consumption, based on Information of Seoul Arts Center Ticket System", *KCTI*, vol.35, pp. 63-65, 2015.
- [21] S. H. Han, O. D. Oh, J. H. Park, J. K. Lee, S. W. Lee, "Analysis of the Social Network Services based Performing Arts Desire Value for the Musical Audiences", *KCTI* vol.40, p.70, 2016.
- [22] J. H. Kim, H. S. Yong, Arts Management, pp. 133-135, Gimm-Young Publishers, 2002.
- [23] McKinsey Global Institute, Big Data: The Next Frontier for Innovation, Competition and Productivity, pp. 78-82, McKinsey Global Institute, 2012.
- [24] Y. J. Jang, New Paradigm for Big Data & Business, p. 42, DIGIECO, 2012.
- [25] B. S. Park, "Actual use of performance information through SNS", *The Musical*, May, 2016, pp. 45-46, 2016.
- [26] W Gordon, R Langmaid. Qualitative Market Research. p.29, AshgatePubCo, 1988.

이 의 신(Eui-Shin Yi)

[정회원]



- 2003년 2월 : 한국예술종합학교 전문사 (예술경영전공)
- 2016년 3월 ~ 현재 : 한국문화경제학회 이사
- 2015년 3월 ~ 현재 : 대한적십자사 자문위원
- 2016년 9월 ~ 현재 : 지역대표공연예술제평가위원
- 2012년 1월 ~ 현재 : 서울사이버대학교 문화예술경영학과 교수

<관심분야>

예술경영, 극장경영, 예술축제, 예술복지

김 선 영(Sun-Young Kim)

[정회원]



- 2015년 2월 : 중앙대학교 예술대학원 문화콘텐츠학과 (문화콘텐츠학 석사)
- 2017년 8월 : 경희대학교 일반대학원 공연예술학과 (예술경영학 박사)
- 2012년 4월 ~ 2014년 3월 : 경기콘텐츠진흥원 산업본부장
- 2015년 3월 ~ 현재 : 예술경영지원센터 대표

<관심분야>

예술경영, 문화콘텐츠, 예술산업