

# 시계열 군집분석을 통한 디지털 음원의 순위 변화 패턴 분류\*

유인진

국민대학교  
비즈니스 IT 전문대학원  
(injin0924@gmail.com)

박도형

국민대학교 경영정보학부/  
비즈니스 IT 전문대학원  
(dohyungpark@kookmin.ac.kr)

본 연구는 현대 사회에서 가장 가치 있는 문화자산이자 한류의 흐름에서 특히 중요한 위치를 차지하는 디지털 음악에 초점을 두었다. 디지털 음악에 대하여 공신력 있는 음원 차트인 ‘가온 차트’에 진입한 음원들의 73주간 순위 변화를 수집하였으며 유사한 특징을 가지는 패턴들로 분류하였다. 이후 각 순위 변화 패턴으로부터 주목할 만한 특징에 대한 설명적 분석을 수행하였다. 구체적으로 음원에 대한 신뢰도 이슈가 발생하기 이전 기간의 국내 발매된 디지털 음원들로 한정하여 시점을 일치시킨 후 시계열 군집분석을 통해 패턴을 도출하고자 하였다. 데이터 수집과 전처리를 통하여 742건의 중복되지 않는 음원들을 확보하였고, 시계열 순위 변화에 대한 시계열 군집분석 결과 16개의 패턴들이 도출되었다. 이후 도출된 패턴들을 기반으로 ‘스테디셀러’와 ‘원 히트 원더’의 두 가지 유형의 대표적인 패턴을 확인하였다. 나아가 두 패턴에 대하여 차트 내에서 음원의 생존 기간과 음원 순위에 관점에서 다섯 가지의 세분화된 패턴으로 분류하였다. 각 패턴들이 가지는 중요한 특징들은 다음과 같다. 원 히트 원더형 패턴에서 아티스트의 슈퍼스타 효과와 편승효과가 강하게 나타났으며, 소비자들의 디지털 음원 선택에 강한 영향을 미친다는 것을 확인하였다. 나아가 스테디셀러형 패턴을 통해서 매우 오랜 시간 소비자들의 선택을 받는 음원들을 확인하였고, 소비자의 니즈를 관통하며 가장 많은 선택을 받는 음원들이 오히려 원 히트 원더형 패턴이 아니라 스테디셀러: 중기 패턴에 포진하고 있음을 확인하였다. 특히 주목할 만한 점은 스테디셀러형 패턴을 통해 기존의 패턴과는 상반되는 ‘차트 역주행’ 현상을 확인했다는 것이다. 본 연구는 디지털 음원을 중심으로 상대적으로 소외되었던 분야인 시간의 흐름에 따른 음원의 순위 변화에 초점을 두었고, 음원의 흥행과 순위를 예측하는 것이 아니라 순위 변화의 패턴을 세분화함으로써 음원 연구에 대한 새로운 접근을 시도하였다는 점에서 의의가 있다.

**주제어** : 음악, 음원, 디지털 음원, 순위 변화, 시계열 군집분석

논문접수일 : 2020년 7월 14일    논문수정일 : 2020년 9월 25일    게재확정일 : 2020년 9월 26일

원고유형 : 일반논문    교신저자 : 박도형

## 1. Introduction

현재 K-POP을 중심으로 가히 모든 문화영역에 있어서 한류에 대한 관심이 증폭되고 있다. 한류는 그 태동기인 “한류 1.0”에서 심화기인 “한류 3.0”로 진화해 왔으며, 그 핵심 장르 또한

과거 드라마로부터 현재 대중문화의 전 영역을 아우르고 있다. 즉 한류의 패러다임이 전환되었다는 것이다. 과거 한류 콘텐츠의 접촉경로가 주로 TV를 통한 접촉이 다수였던 것과는 달리, 현재는 무료 온라인 및 모바일 스트리밍을 통한 접촉이 폭발적으로 증가하였다(Overseas Korean

\* 이 논문은 2020년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2020R1A2C1006001)

Wave Survey Report, 2018). 기존의 드라마 및 TV를 기반으로 흐르던 한류가 ICT의 보급과 SNS의 확산에 따라서 이른바 ‘디지털 한류’로 변화한 것이다.

이러한 현재 한류의 흐름에서는 특히 디지털 음악이 중요한 위치를 차지한다. 이때 디지털 음악은 디지털화되어 유통되는 파일 형식의 음악을 의미한다. 디지털화, 그리고 파일 형식이라는 특성상 디지털 음악은 아날로그 음악과는 달리 네트워크를 통하여 다양한 포맷으로 변환될 수 있으며, PC, 스마트폰, 태블릿 등의 기기의 한계를 두지 않는다. 그리고 언제 어디서든 네트워크에 접속할 수 있는 현대 환경에서 사용자는 사실상 모든 디바이스를 통하여 다운로드 및 스트리밍의 형태로 음악을 감상할 수 있게 되었다.

오늘날의 디지털 음악에 관한 다수의 연구들은 이해관계자들과 음악의 특성을 중심으로 음악이 출시 이후 흥행 성과를 예측하는 연구가 집중적으로 수행되고 있다. 특히 음악의 발매 초기 흥행에 관한 연구가 이목을 끌었는데 이는 이미 스타성이 입증된 아티스트가 슈퍼스타 효과(Superstar Effect)와 편승효과(Bandwagon Effect)로 인하여 더 많은 소비자들에게 노출되고, 더 큰 흥행 성과를 얻을 수 있다는 것이다 (Strobl and Tucker, 2000; Bhattacharjee et al., 2007; Cheom, 2018)

이와 같은 연구는 패러다임의 변화에서 오는 충격을 완화하고, 음악 시장에서 수익을 얻기 위해서는 필수 불가결하다. 그러나 이미 디지털 음악은 생활속에 녹아 들었으며, 아직까지도 흥행 음악 연구(Hit Song Science)는 하나의 풀리지 않는 연구 과제로서 음악 자체의 특성만을 고려하여 흥행 음악(Hit Song)을 예측하는 것은

어렵다는 연구가 다수 제기되고 있다(Pachet and Roy, 2008; Pachet, 2011). 즉, 이제는 연구의 초점을 흥행과 순위 예측에만 두는 것에서 탈피할 필요성이 있다는 것이다. 비록 일부 연구에서 음악의 순위를 기반으로 시간의 흐름에 따른 순위 변화와 시장 경쟁력을 예측하는 연구가 존재하지만, 비교적 짧은 기간의 순위 정보를 다루며 한정된 패턴만을 도출했다는 한계가 있다 (Lee et al., 2014)

본 연구는 음악의 흥행과 순위를 예측하는 것이 아니라 보다 새로운 관점에서 디지털 음악에 대해 접근하려 한다. 먼저 탐색적 분석을 통하여 시간의 흐름에 따른 디지털 음악의 순위 변화를 패턴으로 나타내고 세분화할 것이다. 이후 설명적 분석을 통하여 대표적인 순위 변화 패턴들의 특징을 확인하고자 한다. 본 연구를 통하여 도출된 다양한 유형의 패턴들과 그 특징은 소비자들의 음악 선호와 디지털 음악 그 자체에 대하여 새로운 관점의 시사점을 제공할 수 있을 것이며, 나아가 음악 흥행 및 순위 예측과 같은 기존의 연구 분야에 대해서도 의미 있는 인사이트를 제시할 수 있을 것으로 기대된다.

## 2. Theoretical Background

### 2.1. Digital Music and Consumer

음악은 현대 사회에서 가장 가치 있는 무형 문화자산 중 하나이다. 아티스트들은 음악에 새로운 기술을 적용하기 위해 끊임없이 혁신하고 투자하고 있으며, 음악은 문화산업의 성장을 견인하는 주체로 부상하였다. 특히 디지털 음악의 개념이 대두함에 따라 세계 음악 시장은 빠른

속도로 변화해왔으며, 국내 음원 시장 또한 이러한 흐름에 따르기 위하여 대응하고 있다.

앞서 서술한 바와 같이 디지털 음악은 디지털 화되어 유통되는 파일 형식의 음악을 의미한다 (Kim, 2002). 과거의 아날로그 음악은 LP로부터 CD에 이르는 실물 음악과 대응되는 재생기기, 그리고 MP3 파일을 저장하여 재생할 수 있는 기기에서만 감상이 가능했다. 그러나 디지털 음악은 그 특성상 네트워크를 통하여 다양한 형식으로 변환되며 PC, 스마트폰, 태블릿 등의 기기의 한계를 벗어나기에 이르렀다. 이와 같은 배경에는 모바일 및 네트워크 기술의 발전이 있으며, 이제 소비자는 언제 어디서든 네트워크에 접속하여 가용한 모든 디바이스를 통하여 다운로드 및 스트리밍의 형태로 음악을 감상할 수 있게 되었다.

디지털 음악과 관련된 미디어 채널은 콘텐츠의 유통 및 확산에 결정적인 역할을 하는데, 이는 미디어의 발달로 언제 어디서든 소비자가 원하는 음악을 들을 수 있기 때문이다. 그리고 매체 중에서는 YouTube와 같은 소셜 미디어 서비스와 Facebook, Instagram과 같은 소셜 네트워크 서비스가 일차적이고도 핵심적인 역할을 하고 있다(Jo and Sim, 2013).

나아가 YouTube와 SNS 미디어는 소비자가 디지털 음악을 향유하는데 있어서 중요한 교두보 역할을 하며 시너지 효과를 얻고 있다(Jo, 2013). 모바일 환경이 보편화됨에 따라 접근의 시공간적 제약이 사라진 미디어는 단순한 일방향성 광고에 한정되지 않고, 광고와 음악을 공유하는 소비자들의 개성과 자기표현이 가능한 플랫폼으로 성장하게 되었다. 이에 소비자들의 자발적인 참여와 반응을 토대로 음악 시장의 규모와 질, 다양성을 확장하는데 기여하며 수 많

은 경제적, 문화적 효과를 얻을 수 있게 되었다 (Lee, 2019).

이미 디지털 음악의 판매량은 오프라인 음악을 넘어섰으며, 지속적으로 상승세를 유지하고 있다. 또한 2013년도 이래로 디지털 음악의 스트리밍 수는 1천억을 넘어서며 폭발적인 성장세를 보이고 있다. 즉 사람들의 음악 소비 구조가 오프라인에서 온라인으로 이동했다는 것이다. 또한 음악이 오프라인 중심의 소비에서 온라인 중심으로 변화함에 따라서 소비자가 음악을 향유할 수 있는 진입장벽이 사라졌다고도 말할 수 있다.

이처럼 음악 소비의 주류가 된 디지털 음악의 소비 패턴은 다시 한 번 변화를 맞이하는데, 바로 음악 스트리밍 서비스가 그것이다. 소비 패턴의 주축이 음원 다운로드 서비스에서 스트리밍 서비스로 이동하면서 아티스트들의 음원 발매와 홍보에도 변화가 나타났다. 음원 서비스에서 제공하는 ‘음원 차트’를 중심으로 소비가 집중되기 시작한 것이다. 즉 매일 새로운 음원이 쏟아져 나오는 디지털 환경에서 음원 차트는 음원의 흥행을 좌우하는 중요한 요소가 된 것이다.

그렇다면 스트리밍과 차트가 소비 패턴에 영향을 미칠 수 있는 이유는 무엇인가? 이는 소비자에게 음악이 “경험재”로서 소비되기 때문이다. 소비자는 음악을 직접 감상하기 전까지는 그 특성을 파악할 수 없기 때문에 과거의 감상 경험이나 광고, 방송 등을 통한 간접적 경험을 통하여 소비 결정을 내린다(Kim, 1997). 특히 디지털 음악에 있어서 스트리밍 서비스와 음원 차트는 음악의 경험재적 특성과 능동적 소비자의 역할을 강화한다. 이는 기존의 아날로그 음악이 광고 또는 방송 등의 일방향적 정보전달을 소비

자가 수용하여 감상 및 구매 의사결정을 내렸던 반면, 디지털 음악은 소비자가 스스로 원하는 음악을 원하는 시간에 선택하여 다운로드 및 스트리밍 가능하기 때문이다(Jung, 2016). 이는 과거 수동적 성향의 소비자들이 능동적 소비자로 변화하였음을 시사하고 있으며, 이는 “프로슈머(Prosumer)”의 개념과도 그 맥락을 같이하고 있다.

## 2.2. Paradigm Shift of Music Performance and Consumption Pattern

디지털 음악에 관한 연구는 개념이 대두되었던 시기로부터 확산에 이르기까지 다양한 관점에서 수행되었다. 도입 시기에서는 과거 아날로그 음악에서 디지털 음악으로 변모해가는 음악 시장의 환경 속에서 불법 음원 유통, 인력 불균형, 디지털 음악의 수익분배 등의 다양한 파생 문제들이 나타날 것을 짚어내거나(Moon, 2011), 음악 제작사 및 아티스트의 입장에서 디지털 음악의 대두와 소비 패러다임의 변화 속에서 이해관계자들이 서로 상생할 수 있는 변화의 방향을 제시하기도 하였다(Ryu, 2014).

음악과 이해관계자들 간의 관계에 있어서는 과거와 오늘의 연구 간의 상반된 견해가 나타나기도 한다. 과거 Adler(1985)는 높은 음원 판매 실적을 경험한 아티스트들이 단순히 아티스트들의 재능 이외에도 음원사의 판촉 지원과 같은 외적인 부분이 성공 여부를 결정하는데 중요한 역할을 하는 것을 밝혀냈다. 그러나 디지털 음악 제작과 관련된 이해관계자에 대한 연구에서는 음악 흥행에 있어서 오히려 제작사의 역할은 유의미한 영향을 미치지 못하지만, 음원 유통사들이 흥행에 주요한 위치를 차지한다는 것을 볼

수 있다(Hyeon, 2014).

그렇다면 음악의 흥행에 있어서 어떠한 요인이 중요하게 작용하는가? Strobl and Tucker(2000)는 영화에서 쉽게 찾아볼 수 있는 블록버스터 효과가 음악 시장에서도 유사하게 작용한다는 것을 제시하였다. 이는 음악이 영화 사운드에 사용되었는지가 해당 음악의 차트 순위 유지에 통계적으로 유의한 영향을 미친다는 것이다. 또한 아티스트의 과거 인기도 해당 음악에 매우 강한 영향력을 가지고 있다고 주장하였다. 이른바 슈퍼스타 효과(Superstar Effect)에 의하여 이미 스타성이 입증되거나, 소비자들 사이에서 대중적인 아티스트들이 순위 차트에 더 오래 머물러 있다는 것이다. 물론 해당 연구에서 주장하는 효과가 디지털 음악의 확산 이전에 발견된 것이나 일부 연구를 근거로 디지털 음악 시장에서도 여전히 슈퍼스타 효과가 존재한다는 것을 볼 수 있다(Bhattacharjee et al., 2007). 더욱이, 유행과 트렌드 변화에 민감한 음악 소비는 편승효과(Bandwagon Effect)가 강하게 나타내기 때문에, 미디어를 통하여 공유되는 소비자의 반응이 음악 소비에 민감하게 작용한다.

그러나 여전히 흥행 음악 연구(Hit Song Science)는 하나의 풀리지 않는 연구 과제로서 음악 자체의 특성만을 고려하여 흥행 음악(Hit Song)을 예상하는 것은 어렵다는 연구가 다수 제기되고 있다(Pachet and Roy, 2008; Pachet, 2011). 따라서 이제는 음악 연구에 대한 접근 방식에 변화가 필요하다는 것이다. 그렇다면 새로운 유형의 접근 방식은 무엇이 될 수 있는가? 앞서 음악은 소비자에게 경험재로서 소비된다고 하였다. 이와 같은 특성은 디지털 음악에서 스트리밍 서비스와 음원 차트를 통하여 강화된다. 그리고 프로슈머의 개념과도 그 맥락을 같이한다고 하였

다. 그리고 이는 기존의 연구의 미디어를 통하여 공유되는 소비자의 반응이 음악 소비에 민감하게 작용한다는 주장을 뒷받침한다. 따라서 음악의 경험재적 특징을 고려한 접근이 그 대안이 될 수 있을 것이다.

음악의 장르에서도 이러한 경험재로서의 특징을 볼 수 있는데, 이미 소비자에게 익숙한 장르의 음악이 아니라 들어보지 못한 새로운 유형의 장르 음악들이 소비자의 선택을 받을 수 있기 때문이다(Han, 2004). 그리고 이미 음악은 그 자체로 브랜딩 효과를 보이고 있기 때문에, 음악을 소비자에게 밀접하게 다가갈 수 있도록 하는 것이 음악을 생산한 아티스트의 브랜드 가치를 높일 수 있다는 의견 또한 존재한다(Lee, 2019). 특히 음악과 소비자 중심의 연구에서는 경험재적 특징이 디지털 음악에 있어서 어떠한 성격으로 변화해 나아가는지 보이기도 하였다(Jung, 2016).

한편 음악 스트리밍 서비스를 기점으로 디지털 음악의 소비 패턴이 변화함에 따라 새로운 문제가 야기되기도 하였다. 매시간 수 많은 아티스트들로부터 쏟아져 나오는 음악들 속에서 소비자들은 과거 정보의 홍수같은 ‘음악의 홍수’ 환경에 처해진 것이다. 즉, 감상할 수 있는 음악은 매우 방대하지만, 그 속에서 어떠한 음악을 감상할지 선택하는 것에서 어려움을 느끼게 된 것이다. 그러나 이와 같은 소비자의 어려움은 음원 서비스에서 제공하는 음원 차트로 소비가 집중되며 어느정도 완화되었다. 소비자들이 음원 차트를 통하여 어느정도 검증된, 그리고 유행과 트렌드를 이끌어가는 음악들을 선별하여 소비할 수 있게 된 것이다.

이처럼 음원 차트는 음악 소비의 중심이 되면서 디지털 음악의 흥행을 좌우하는 주요 요인으

로 자리잡았으며, 음악 연구의 새로운 도메인으로 부상하게 되었다. 이때 비교적 최근에 등장한 음원 차트와 관련된 현상에 있어서 주목할 만한 것이 있다. 이른바, ‘차트 역주행’으로 대표되는 순위 변화 패턴은 기존의 음원이 처음 공개되었을 때 소비자들이 해당 곡에 가장 관심을 가지고, 그 후 자연스럽게 관심이 식어가는 기존의 패턴과는 상반된다. 즉, 처음 출시 혹은 출시 이후에도 긴 시간 순위에 들지 못하거나, 낮은 순위를 차지하고 있음에도 불구하고, 시간이 흐름에 따라 혹은 특정 사건을 기점으로 그 순위가 역전하여 음원 차트의 높은 위치에 안착한다는 것이다.

그러나 차트 역주행과 같이 기존의 음악 소비 패턴과는 상반된 현상이 대두되었음에도 불구하고 관련된 연구는 미흡하다는 한계가 있다. 비록 일부 연구에서 음원의 순위를 기반으로 시간의 흐름에 따른 순위 변화와 시장 경쟁력을 예측하는 연구가 존재하지만, 단기간의 순위 정보를 다루며 한정된 패턴만을 도출하였다(Lee et al., 2014). 비교적 짧은 60일 간의 순위 데이터를 사용하여, K-means를 통하여 차트 내 음악들을 완만유지형, 급격하강형의 두 그룹으로 분류한 것이다. 따라서 음원 차트로부터 보다 다양한 소비 형태를 보이는 패턴들을 찾아내고, 패턴들의 특징을 규명할 수 있다면 음악 연구에 있어 새로운 관점의 시사점을 제공할 수 있을 것으로 기대된다.

### 2.3. Time-Series Clustering

다양한 소비자들 속에서 동질의 소비자들을 분류해내는 것과 그 소비자 집단의 특성을 파악하는 것은 군집분석에 있어 핵심이라 할 수 있

다. 소비자에 대해서 동질의 집단이라는 것은 다른 집단의 소비자들과는 이질적이면서도, 같은 집단 내 소비자 집단에서는 보편적인 것을 의미한다. 따라서 동질 집단 속의 소비자를 통하여 다른 소비자들을 합리적으로 추론해 볼 수 있고, 그 소비 유형 또한 예측 가능하다고 할 수 있다. 본 연구는 음원 차트의 시계열 데이터에 대한 군집분석(Time-Series Clustering) 방법론을 통하여 음원 순위에 따른 패턴 분류를 수행하고자 한다.

군집 분석은 표본들 간의 유사도에 기반하여 동질성이 높은 표본들을 하나의 군집으로 분류하는 탐색적 분석을 의미한다(Wilks, 2011). 데이터 마이닝 기법들 중 대표적인 비지도학습인 군집 분석은 분석 과정에서 동질적인 표본끼리는 동일한 군집에 배치하고, 이질적인 표본들 간에는 서로 다른 군집에 속하도록 배치한다. 따라서 분석의 결과로 구분된 각 군집들은 집단 내에서는 높은 동질성과 유사성이 나타나게 된다(Berry and Linoff, 1997). 군집 분석의 주요 목적은 데이터들 간의 구조적 특성을 파악하기 위하여 주로 사용된다. 즉, 군집화를 통하여 복잡한 데이터들 몇 개의 군집으로 나타낼 수 있으며, 나아가 데이터 집단의 대표적인 특성이나 유형을 파악하기 용이하다는 장점이 있다(Bishop, 2006).

군집 분석은 유사한 데이터들을 동질한 군집으로 분류하는 것이 중요하다. 따라서 데이터들 간의 유사성이나 비유사성을 정량화 한 거리를 기반으로 군집화가 수행되며, 데이터의 성격이나 분석의 목적에 따라서 다양한 형태의 거리 척도가 존재한다. 이처럼 거리는 군집 분석에서 매우 중요한 핵심이지만, 시계열 데이터는 시간의 흐름에 따른 동적 특성으로 인하여 유사성에

따른 거리 개념을 적용하기 어려운 점이 있다. 특히 기존 군집분석에서 사용되는 유사성은 시계열 데이터 값 사이의 상호 의존 관계를 고려하지 않기 때문에 바람직하지 않다. 따라서 시계열 데이터 간의 유사성 및 비유사성을 정의하여 군집분석을 수행하기 위한 다양한 시도가 제안되어 왔다.

특히, 동적 시간 워핑(DTW; Dynamic Time Warping)과 Global Alignment Kernel(GAK)이 다수의 연구에서 사용되고 있다. 두 기법은 시계열 데이터가 가지는 패턴의 유사성을 측정하는 알고리즘으로, 관측시점에 상관없이 두 계열의 신호를 비교하는데 주어진 두 계열 사이의 최적적합을 계산한다. 이는 동일한 관측시점 하에서의 유클리드 거리에 기초한 기존의 군집분석과는 다르다. 구체적으로 한 계열의 관측시점을 변경하면서 다른 계열의 임의 시점에 대한 최소 거리를 계산하여 이를 바탕으로 두 계열 간의 유사성 정도를 측정하게 된다. 이 방법의 또 다른 특징은 두 계열의 관측시점이 동일하지 않아도 측정할 수 있을 뿐만 아니라 표본크기가 동일하지 않아도 계산할 수 있다는 점이다. 따라서 본 연구에서는 유사성 척도로 동적 시간 워핑과 GAK을 고려하였으며, 둘 중 상대적으로 개선된 방법론인 Global Alignment Kernel(GAK)을 사용하였다(Cuturi et al., 2007).

### 2.3.1. Global Alignment Kernel (GAK)

Global Alignment Kernel은 Cuturi et al.(2007)이 제시한 방법론으로 DTW보다 빠르고 효율적이라는 특징을 가진다. GAK 방법론의 커널은 데이터의 차원을 고차원/저차원으로 매핑하는 기법으로 SVM과 같은 머신 러닝 알고리즘에서

주로 사용된다. 즉, GAK은 커널을 통하여 선형 분리가 불가능하게 섞여 있는 시계열 데이터를 선형 분리가 가능한 고차원으로 매핑 후, 두 군집을 분류하는 초평면을 찾고, 다시 저차원으로 매핑해 시계열 데이터 간의 유사성을 계산할 수 있다. 따라서 구조화된 데이터에 대하여 커널을 활용하면, SVM과 같이 분석을 단순화할 수 있어 속도와 효율성을 높일 수 있다.

### 2.3.2. Partitioning Around Medoid (PAM)

Partitioning Around Medoid에서 Medoid는 한 군집에 포함된 표본을 나타내며, 해당 표본이 속한 군집 내의 다른 모든 표본과의 유사도 또는 거리가 가장 작은 개체를 말한다. 즉, 한 군집의 중심점에 있는 표본이라고 할 수 있다. PAM은 평균을 중심으로 군집화를 수행하는 방법보다 효율적인 측면이 있다. 가령, Medoid를 사용하는 대표적인 방법인 K-medoid의 경우, 평균을 중심으로 군집화를 수행하는 K-mean에 비해 이상치나 데이터의 노이즈에 대하여 상대적으로 강건한 특성이 있다. 이는 평균을 중심으로 계산하는 것이 이상치의 영향을 크게 받는 반면, PAM은 중심점에 있는 표본을 사용하여 대응을 통해 군집화를 수행하기 때문이다.

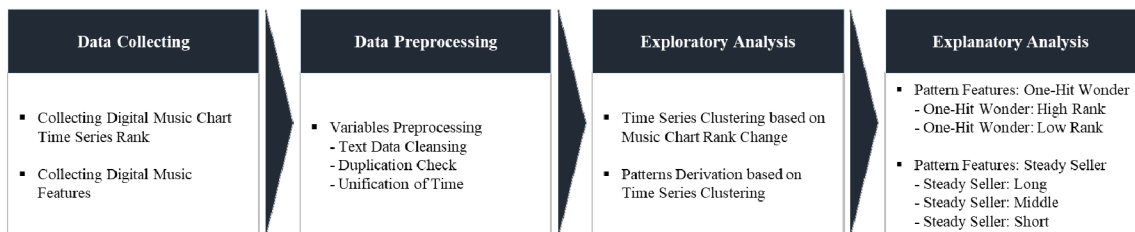
## 3. Research Model

### 3.1. Research Process

전술한 내용에 기반하여 본 연구가 제시하는 연구 프로세스는 다음과 같다. 먼저 데이터 수집 과정에서는 디지털 음원들의 순위 변화를 확인하기 위하여, 각 음원의 시계열 데이터를 수집하였다. 이후 데이터 가공 단계에서는 수집된 데이터에 대하여 음원의 시간의 흐름에 따른 순위를 매칭하고, 곡명과 아티스트명을 중심으로 처리하였다. 이후 각 분석은 탐색적 분석(Exploratory Analysis)과 설명적 분석(Explanatory Analysis)으로 구성된 두 단계가 순차적으로 수행된다. 첫 번째, 탐색적 분석 단계에서는 시계열 군집분석을 통하여 음원의 순위 변화를 시각화하고, 대표적인 패턴과 세분화된 패턴들을 도출하였다. 두 번째, 설명적 분석 단계에서는 각 패턴이 구체적으로 어떠한 특징을 가지는지, 그리고 그 원인은 무엇으로부터 기인하며, 슈퍼스타 효과, 편승효과, 그리고 차트 역주행 현상을 확인할 수 있는지 설명하고자 하였다. 구체적인 연구 모형은 다음과 같다.

### 3.2. Public Confidential Music Chart

본 연구는 음원의 순위 변화를 시각화하고,



(Figure 1) Research Process

대표적인 패턴을 도출하고자 한다. 이때 데이터의 출처에 대하여 다소 조심스러운 접근이 필요한데, 그 이유는 다음과 같다. 2019년을 기점으로 음원 사재기가 음악 시장의 이슈로 대두되었고, 이는 음원 서비스 플랫폼의 공정성과 관련한 문제들이 수면위로 부상하는 계기가 되었다. 구체적으로 음악적 인지도가 낮은 아티스트가 발매한 음원이 일부 음원 플랫폼에서 재생 회수가 상대적으로 낮은 새벽 시간에 기형적으로 높은 재생 회수를 보이며 차트 순위에서 1위에 오른 것이다. 이에 대하여 타 아티스트들 및 소비자들은 해당 음원과 아티스트에 대하여 음원 사재기에 대한 의혹을 제기했지만, 해당 이해관계자들은 이를 부인하고 있다.

그러나 사실 여부와는 무관하게 이전부터 음원 서비스 플랫폼에 대한 불합리성과 아티스트들의 부당한 처우에 대한 개선 요구는 지속적으로 대두되었다. 이에 국내 음악 시장의 발전과 부당함을 소거하고자 음악적 공정성을 담보할 수 있는 공인 차트가 필요성이 제시되었다. 그리고 문화체육관광부가 제작사와 아티스트들의 의견을 수용하여 공정성이 담보된 ‘가온 차트’가 등장하였다.

따라서 본 연구는 음원 차트의 공정성을 확보하고자 음원에 대한 순위정보 데이터를 ‘가온 차트’를 통하여 수집하였다. 가온 차트는 한국음악콘텐츠협회(문화체육관광부 주관 사단법인)에서 미국의 빌보드 차트 및 일본의 오리콘 차트와 같이 국내에서도 공인 음악 차트의 필요성이 대두됨에 따라 음원의 순위와 정보를 제공하고 있는 서비스이다. 가온 차트는 국내에서 서비스중인 사설 음원 차트와는 달리 정부 기관에서 공인한 차트로 기본적인 신뢰성이 확보되며, 타 서비스에서 제공되는 순위 정보보다 공

신력이 있다고 볼 수 있다.

### 3.3. Data Collection and Preprocessing

본 연구는 분석에 활용하는 음원에 대한 순위 정보 데이터를 가온 차트로부터 수집하여 데이터의 공정성과 신뢰성을 확보하고자 하였다. 구체적으로 가온 차트로부터 제공되는 다양한 유형의 음원 데이터들 중 연구 목적에 부합하도록 ‘디지털 음원의 주간 순위’ 데이터를 수집하였다. 데이터의 수집 기간은 2017년 12월 31일부터 2018년 01월 06일을 첫 주차로 시작하여 2019년 05월 19일부터 2019년 05월 25일까지의 73주이다. 수집된 데이터의 내용은 다음과 같다. 수집 기간 내에서 음원 차트상 상위 100위권 내에 진입한 음원을 대상으로 그 기간과 순위, 음원 명, 아티스트 명, 앨범 명, 가온지수, 제작사, 유통사에 대한 데이터를 수집하였다.

구체적인 수집 및 데이터 전처리 과정은 다음과 같다. 데이터 수집을 수행하기 위하여 R 소프트웨어를 사용하였으며, 원활한 수집을 위하여 URL을 세분화하여 각 기간별 순위정보를 제공하는 URL에 반복 접근하도록 하였다. 이후 ‘xml2’ 패키지를 통하여 웹 페이지의 소스코드 내에 텍스트 형태로 작성된 음원 정보를 테이블 형식으로 저장하였다. 이후 테이블 내에서 분석에 필요한 데이터를 선택하고, 텍스트 내 가시성을 저해하는 공백문자 및 특수문자를 소거한 후 음원 명, 아티스트 명, 앨범 명으로 구분하였다. 마지막으로 음원 또는 아티스트의 이름이 동일한 케이스를 모두 구분하기 위하여 ‘음원 이름 by 아티스트 이름’의 구조를 가지는 텍스트 형식의 고유 키를 구성하였다.

앞선 데이터 수집을 통하여 총 73주 간 음원



<Table 1> Time series clustering results

No.	Intra-cluster distance	Size	No.	Intra-cluster distance	Size
1	0.000029301	53	9	0.000002251	64
2	0.000030519	31	10	0.000113201	53
3	0.021334190	29	11	0.000460156	43
4	0.000102907	34	12	0.000647854	45
5	0.000000886	29	13	0.000004288	49
6	0.002865161	55	14	0.001380504	58
7	0.000210031	55	15	0.000008915	44
8	0.007082197	72	16	0.003018437	28

차트 상위 100위 내에 포함된 7,300건의 음원이 확인되었다. 한편, 디지털 음원의 경우 2주 이상 음원 차트에 포함되는 케이스가 빈번하기 때문에 전처리 과정을 통하여 음원의 중복을 제거하고 시점을 통일하였다. 중복되는 음원에 대해서는 중복 확인 함수를 통하여 그 수와 위치를 확인한 후, 소거하여 분석을 위한 데이터를 구성하였다. 이를 통하여 사전 7,300건의 음원 데이터들 중 분석을 위한 최종 742건의 고유한 음원 리스트가 확보되었으며, 마지막으로 각 음원이 차트에 등재된 시점이 다르기 때문에 음원의 발매 시점을 모두 통일하여 데이터를 구성하였다.

## 4. Research Results

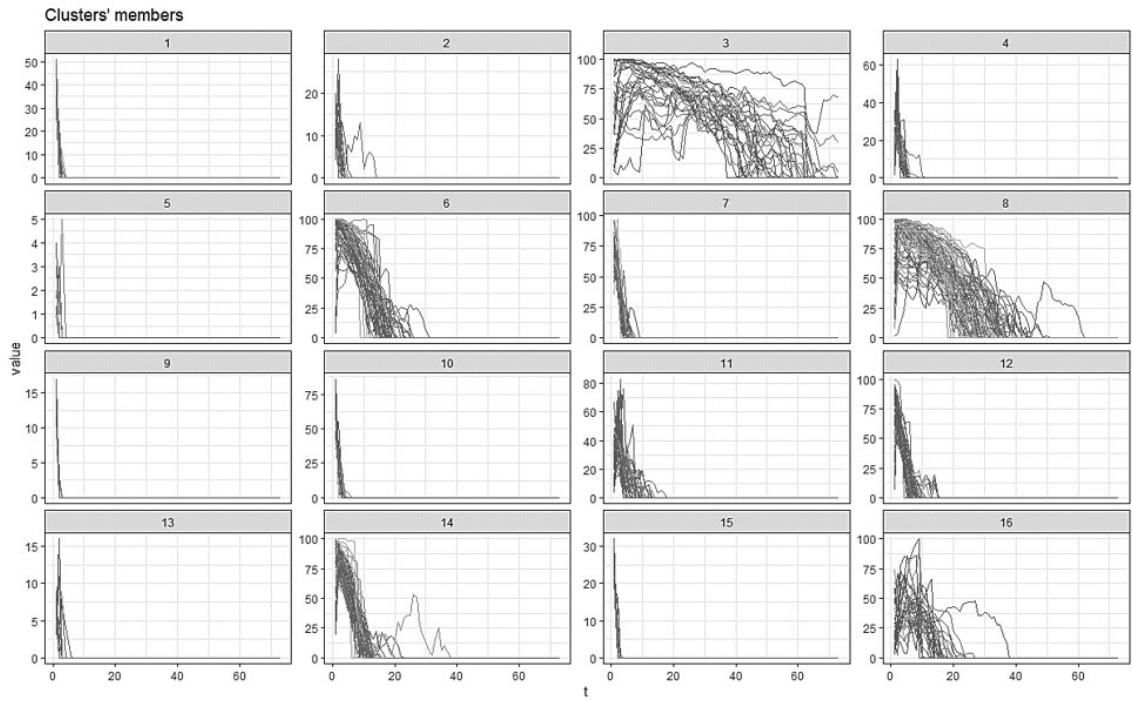
### 4.1. Pattern Derivation based on Time Series Clustering

본 연구에서는 시계열 군집분석을 수행하기 위하여 R 소프트웨어에서 제공하는 ‘tsclust’ 패키지를 사용하였다. 해당 패키지는 시계열 데이터에 대한 군집분석을 수행할 때 최적화되어 있기 때문에 본 연구에서 활용하려는 군집화 방법

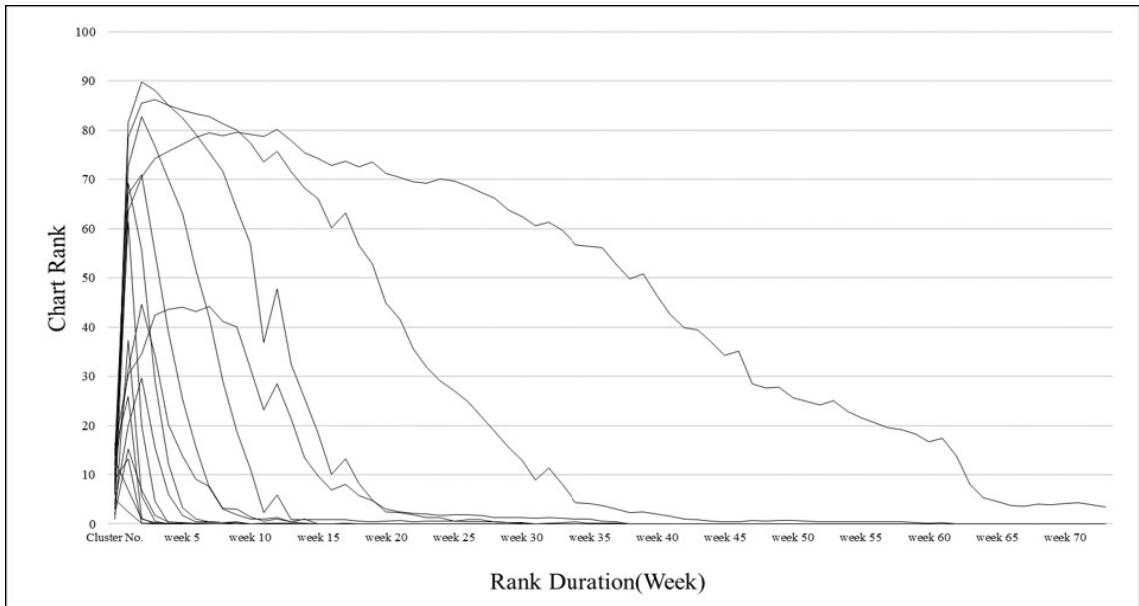
론을 적용하기 용이하다(Pablo Montero, 2014). 해당 패키지를 통하여 시계열 군집분석을 수행하였으며, 군집화 수행을 위한 Parameter로 유사도 척도에는 Global Alignment Kernel을 중심 추출에는 Partitioning Around Medoid를 사용하였다. 각 음원의 순위 변화에 따른 군집 분석 결과는 <Table 1>과 같다.

742건의 음원들은 순위 변화 패턴에 따라 16개의 군집으로 분류되었다. 각 군집들은 평균적으로 46.375건의 음원들로 구성되며, 평균 군집 내 거리는 0.00233으로 나타났다. 군집분석 결과를 살펴보면, 8번 패턴으로 분류된 음원들이 가장 많았으며(72건), 16번 패턴으로 분류된 음원이 가장 적은 것이 확인되었다(28건). 각 군집에 따라 음원의 순위 변화 패턴을 시각화한 결과는 <Figure 2>와 같으며, 각 군집간 상이한 순위 변화가 나타나는 것을 볼 수 있다.

한편 군집분석 결과와 그래프를 통하여 각 패턴들의 대략적인 특징은 확인할 수 있으나, 구체적인 특징을 이해하기에는 어려움이 있다. 따라서 각 군집의 중심을 계산하여 명확한 순위 변화 패턴을 확인하고, 유사한 패턴 유형들로 분류하고자 한다. 각 군집의 중심을 계산하여



〈Figure 2〉 Time Series Clustering Results



〈Figure 3〉 Cluster' s Centroid

시각화한 결과는 <Figure 3>과 같다.

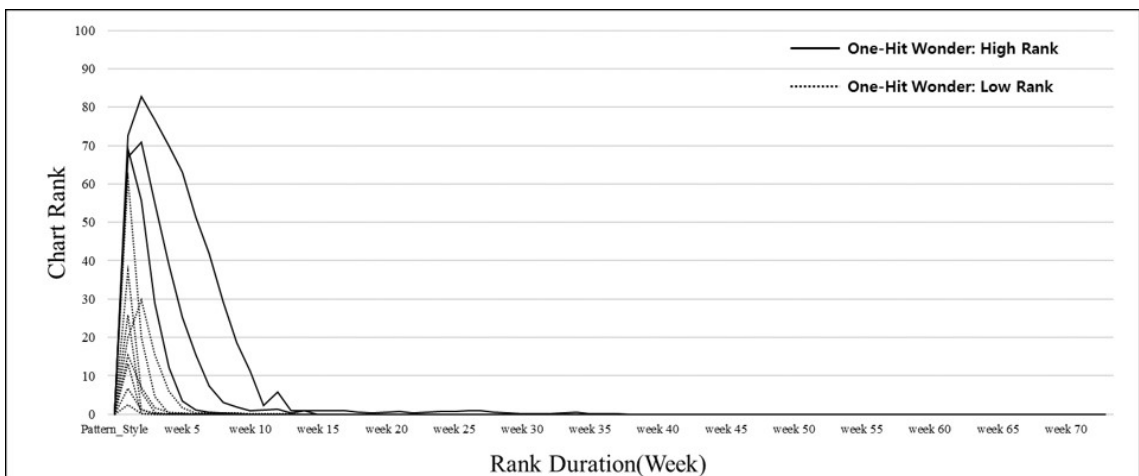
5개의 군집을 제외하면 대다수의 음원들은 10주차가 지나기 전 음원 순위에서 이탈하고 있음을 볼 수 있다. 특히 다수의 군집이 5주 이내에 순위 밖으로 이탈하고 있다. 반면 제외된 5개의 군집들은 짧게는 20주차에서 길게는 전체 데이터 수집 기간인 73주차에 이르기까지 음원 차트 내에 생존하고 있음을 볼 수 있다. 정리하자면 발매 후 순위가 상승하지만 빠르게 이탈하는 패턴을 가지는 군집들에 대해서는 “원 히트 원더(One-Hit Wonder)”형 패턴으로, 상대적으로 오랜 시간 차트 내에 머무르는 패턴을 가지는 5개의 군집에 대해서는 “스테디셀러(Steady Seller)”형 패턴으로 정의할 수 있겠다.

#### 4.2. Pattern Features: One-Hit Wonder

원 히트 원더형 패턴을 나타내는 군집은 1, 2, 4, 5, 7, 9, 10, 12, 13, 14, 15번의 11개 군집으로 음원 발매 혹은 음원 차트에 짧은 시간 내에 빠르게 진입 후, 시간의 흐름에 따라 차트 순위가

급격히 감소하고 있다. 차트에서 빠르게 이탈하며, 순위의 변화폭이 크다는 점을 고려하자면 해당 군집에 포함된 음원들은 음원 발매 이전 이슈성으로 관심을 모으거나, 특정 팬층 및 계절을 겨냥하여 소비자의 이목을 집중한 음원으로 볼 수 있겠다.

한편 원 히트 원더형 패턴에서도 상대적인 생존 기간과 최대 음원 순위를 고려하여 두 가지 유형의 패턴으로 세분화할 수 있다. 최대 음원 순위에 더 초점을 두자면 각 패턴은 고순위와 저순위로 분류할 수 있으며, 다음과 같은 특징을 가진다. 첫 번째, 원 히트 원더: 고순위 패턴의 경우 7, 12, 14번 군집으로 구성되며, 각각 55건, 45건, 그리고 58건으로 총 158건의 음원들이 포함된다. 평균적으로 약 10주차 동안 차트 순위에서 머무르면서, 차트 상위 약 30위권에 도달하며 상대적으로 많은 소비자들의 선택을 받고 있다. 주목할 만한 점은 해당 패턴에 포함된 음원들의 대부분이 ‘방탄소년단’, ‘Wanna One(워너원)’ 그리고 ‘볼빨간사춘기’ 등으로 대표되는



<Figure 4> One-Hit Wonder Centroid

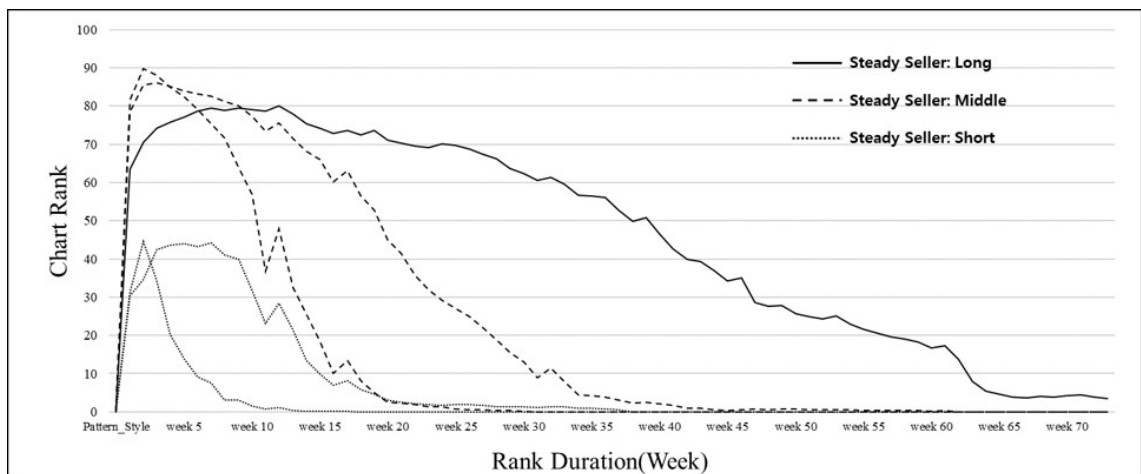
저명한 아티스트들과 ‘쇼미더머니’, ‘고등래퍼’ 등으로 대표되는 대국민 공개 오디션 프로그램으로부터 발매된 음원이라는 것이다. 즉 원 히트 원더: 고순위 패턴은 이미 견고하게 자리잡은 팬층을 보유한 아티스트들의 ‘슈퍼스타 효과’가 디지털 음악 시장에서도 여전히 존재한다는 것을 뒷받침한다. 나아가 공개 오디션 프로그램으로부터 발매된 음원을 통하여 ‘편승효과’에 의한 소비자들의 선택 또한 확인해 볼 수 있다. 구체적으로 프로그램이 높은 시청률을 확보하며 미디어를 통하여 빠르게 공유되고 있으며, 음원 발매 이전 방송 송출을 통하여 소비자들의 관심을 모을 수 있다는 점에서 유행과 트렌드 변화에 민감한 편승효과를 강하게 받는다는 것이다.

두 번째, 원 히트 원더: 저순위 패턴의 경우 1, 2, 4, 5, 9, 10, 13, 15번 군집으로 구성되며, 각각 53건, 31건, 34건, 29건, 64건, 53건, 49건 그리고 44건으로 총 357건의 음원들이 포함된다. 가장 많은 수의 군집과 음원들이 포함되며 평균적으로 약 5주 이내의 매우 짧은 기간 동안 차트 순

위에 머무르면서, 차트 상위 약 60위권 바깥에 포진하여 상대적으로 적은 소비자들의 선택을 받고 있다. 전체 음원들 중 약 48%에 해당하는 음원이 해당 유형으로 분류되며, 음원 시장에서 보편적인 순위 변화 패턴으로 볼 수 있다. 패턴의 특성상 음원 발매 직후 또는 차트 진입 이후 빠르게 이탈하며, 급격한 순위 변화폭을 가지기 때문에 소비자 니즈나 기대를 충족하지 못하여 차트에서 이탈하는 음원으로 해석해 볼 수 있다. 그러나 본 연구의 차트 순위 수집 시점에 긴 시간 동안 차트에 머물렀지만 해당 시점에서는 감소 추세가 나타나는 음원일 수 있음을 고려하여 해석해야 할 것이다.

### 4.3. Pattern Features: Steady Seller

스테디셀러형 패턴을 나타내는 군집은 3, 6, 8, 11, 16번의 5개 군집으로 음원 발매 혹은 음원 차트 진입 이후에 상대적으로 긴 시간동안 순위권 내에 유지되며, 그 순위 또한 시간의 흐름에 따라서 완만하게 변화하고 있다. 차트 내



〈Figure 5〉 Steady Seller Centroid

에 오랜 기간 유지되며, 순위의 변화폭이 크지 않다는 점을 고려하자면 해당 군집에 포함된 음원들은 일시적인 이슈정보보다는 음악성이 더 높기 때문에 음원 소비자들로부터 지속적으로 소비되고 있음을 볼 수 있다.

스테디셀러형 패턴에서도 차트 내 생존 기간이나 최대 음원 순위를 고려하여 세 가지 유형의 패턴으로 세분화해 볼 수 있다. 생존 기간에 더 초점을 두자면 각 패턴은 장기, 중기, 단기로 분류할 수 있으며, 다음과 같은 특징을 가진다. 첫 번째, 스테디셀러: 장기 패턴의 경우 오직 3번 군집 하나로 구성되며, 29건의 음원들이 포함된다. 약 60주차 이상의 매우 긴 시간 동안 소비자의 선택을 받고 있으며, 순위 변화폭도 매우 완만하다. 29건의 음원들 중 “그리워하다 by 비투비(BTOB)”, “밤편지 by 아이유(IU)” 그리고 “봄날 by 방탄소년단”의 세 음원은 전체 73주차 동안 차트 순위에서 이탈하지 않는다. 이는 약 1년 5개월의 긴 시간이라는 한계와 매시간 수많은 아티스트들로부터 쏟아져 나오는 ‘음악의 홍수’ 같은 환경 속에서도 꾸준히 소비자들의 선택을 받고 있다는 것이다.

두 번째, 스테디셀러: 중기 패턴의 경우 6, 8번 군집으로 구성되며, 각각 55건과 72건으로 총 127건의 음원들이 포함된다. 평균적으로 약 30주차 이상의 기간 동안 소비자의 선택을 받고 있지만, 장기 패턴에 비하여 상대적으로 순위 변화폭이 크다. 그러나 주목할 만한 점은 30주라는 시간은 여전히 6개월 이상의 매우 긴 시간이라는 것과 모든 군집들 중에서도 해당 패턴에 속하는 음원들이 가장 높은 순위를 차지한다는 것이다. 즉 소비자의 니즈를 관통하고 가장 많은 선택을 받은 음원들이 스테디셀러: 장기 혹은 원 히트 원더형 패턴이 아니라, 오히려 중기

패턴에 포진하고 있음을 확인할 수 있다.

세 번째, 스테디셀러: 단기 패턴의 경우 11, 16번 군집으로 구성되며, 각각 43건과 28건으로 총 71건의 음원들이 포함된다. 평균적으로 약 15주차 동안 차트 순위에 머무르며 상대적으로 단기간 동안 소비자의 선택을 받고 있다. 한편 다른 스테디셀러 패턴과는 달리 평균 50위권 내에 진입하지 못하는 것을 볼 수 있다. 따라서 다수의 소비자 니즈를 충족하지는 못하지만, 특정 아티스트나 장르 음원에 대한 두터운 팬층으로 인하여 일부 소비자들의 지속적인 선택을 받는 것으로 해석해 볼 수 있다.

한편, 스테디셀러 패턴들 사이에서는 특히 주목할 만한 점을 확인할 수 있다. 이른바, ‘차트 역주행’으로 대표되는 순위 변화 패턴이 확인되는데, 이는 기존의 음원이 처음 공개되었을 때 소비자들이 해당 곡에 가장 관심을 가지고, 그 후 자연스럽게 관심이 식어가는 기존의 패턴과는 상반된다. 구체적으로는 스테디셀러: 장기 패턴에 속하는 3번 군집과 단기 패턴에 속하는 16번 군집의 순위 변화가 그러하다. 각 군집에 속하는 음원들은 발매 초기 혹은 차트 진입 직후에는 상대적으로 낮은 순위권에 포진되어 있는 반면, 각각 12주차 그리고 7주차에 이르러서는 높은 순위권을 차지하며 ‘차트 역주행’ 현상을 반영하고 있다. 차트 역주행 현상을 현실에서 쉽게 접하기 어려운 것을 고려하자면, 각 군집에 포함된 음원의 수가 29건, 28건으로 전체 군집들 중에서 가장 적은 수라는 점 또한 이를 뒷받침하는 근거로 해석해 볼 수 있다. 그러나 두 패턴의 형태는 상대적으로 유사하지만 차트 역주행으로 인해 최고점에 도달하고 안정되는 순위는 서로 상이하다. 따라서 차트 역주행 현상이 예상되는 음원들의 순위 안정화가 이루어

지기 시작하는 정도를 고려하면 해당 음원의 최종 차트 순위와 차트 내 생존 기간을 예측하는데 도움이 될 수 있을 것으로 보인다.

## 5. Conclusion and Implication

본 연구는 현대 사회에서 가장 가치 있는 문화자산이자 한류의 흐름에서 특히 중요한 위치를 차지하는 디지털 음악에 초점을 두었다. 디지털 음악에 대하여 공신력 있는 음원 차트에 진입한 음원들의 순위 변화에 기반하여 유사한 특징을 가지는 패턴들을 분류하고, 각 순위 변화 패턴으로부터 주목할 만한 특징들을 설명하였다. 구체적으로 음원들의 시계열 순위 변화를 기반으로 ‘스테디셀러’와 ‘원 히트 원더’의 두 가지 대표적인 패턴 유형으로 분류하였으며, 각 패턴을 차트 내 생존 기간과 순위에 관점에서 다섯 가지 패턴으로 세분화하였다. 나아가 각 패턴들이 가지는 중요한 특징들을 확인하였다. 그 결과 원 히트 원더형 패턴에서 아티스트의 슈퍼스타 효과와 함께 편승효과가 소비자들의 디지털 음원 선택에 영향을 미친다는 것을 확인하였다. 나아가 스테디셀러형 패턴을 통해서 매우 오랜 시간 소비자들의 선택을 받는 음원들을 확인하였고, 소비자의 니즈를 관통하며 가장 많은 선택을 받는 음원들이 오히려 스테디셀러: 중기 패턴에 포진하고 있음을 확인하였다. 특히 주목할 만한 점은 스테디셀러형 패턴을 통해 기존의 패턴과는 상반되는 ‘차트 역주행’ 현상을 확인했다는 것이다.

본 연구는 다음과 같은 공헌을 갖는다. 이론적 공헌으로 첫 번째, 디지털 음원을 중심으로 음원의 흥행과 순위를 예측하는 것이 아니라 순

위 변화의 패턴을 세분화함으로써 음원 연구에 대한 새로운 접근을 시도하였다. 기존의 디지털 음원 관련 연구는 이해관계자들과 음악 자체의 특성을 중심으로 음원 발매 이후 흥행 성과를 예측하는 연구가 집중적으로 수행되었다. 그러나 본 연구는 음원에 대하여 상대적으로 소외되었던 분야인 시간의 흐름에 따른 음원의 순위 변화에 초점을 두어 연구를 수행하였다. 그리고 이를 통하여 한정된 패턴만을 도출했던 선행 연구에서 확장하여 보다 세분화된 패턴들을 도출하였으며, 음원의 성과 예측 관점에서도 유의미한 인사이트를 제공했다는 점에서 의의가 있다. 두 번째, 음원의 순위 변화 패턴에 기반한 설명적 분석을 통하여 디지털 음원에 관한 다양한 효과와 현상을 근거하였다. 다수의 선행 연구에서 검증되었던 슈퍼스타 효과와 편승효과가 디지털 음원의 순위 변화 패턴을 통해서도 확인할 수 있었으며, 그 효과들을 강하게 받는 음원들의 패턴과 특징 또한 확인하였다. 한편 차트 역주행과 같이 기존의 음악 소비 패턴과는 상반된 현상이 대두되었음에도 불구하고 관련된 연구가 미흡한데, 본 연구를 통하여 차트 역주행이 발생하는 음원들과 그 패턴을 확인하고, 나아가 패턴에 기반하여 차트 순위와 생존 기간 예측에 관한 시사점을 제시하였다.

실무적 공헌으로 첫 번째, 장기간의 시계열 데이터를 체계적인 방법으로 정제하여 유의미한 패턴들을 도출하고, 분석에 활용하는 일련의 프로세스를 제시하였다. 스마트 기술의 발전으로 분야를 가리지 않고 다양한 시계열 정보의 수집과 축적이 이루어지고 있는 현 시점에서 시계열 데이터를 체계적으로 정제하여 유의미한 인사이트를 추출해내는 것이 중요한 과제로 대두되었다. 본 연구는 장기간의 시계열 데이터를

처리하여 음원들의 순위 변화 패턴을 도출하고 설명하였다. 본 연구가 진행한 접근법은 음원이라는 도메인에 한정하지 않고, 사용 로그를 수집하고 있는 산업 환경으로부터 사용자 이용 행태를 수집하고 있는 온라인 환경의 데이터 분석에까지 확장하여 적용 가능하다는 장점이 있다. 두 번째, 음원의 패턴이 가지는 특징들을 통하여 소비자의 니즈를 충족하고 소비자의 선택을 받기 위한 방향성을 제시하였다. 본 연구의 결과를 토대로 각 유형의 패턴에서 소비자의 니즈가 어느 수준으로 충족될 수 있는지 식별할 수 있었다. 따라서 디지털 음원들이 소비자들의 니즈를 충족하고, 소비자의 선택을 받기 위한 패턴의 특징들을 고려한다면 차트 순위와 생존 기간을 개선하는데 도움이 될 수 있을 것이다.

한편 본 연구에서 충분히 고려되지 못한 사항에 대해서는 후속 연구 진행을 통하여 점진적인 보완이 필요할 것으로 판단된다. 첫 번째, 차트 상 제공되는 한정적인 음원 정보 외에 각 음원의 장르 등의 음악적 속성과 아티스트의 성별, 솔로 및 집단 여부 등의 속성은 패턴에 대한 설명적 분석에 포함되지 않았다. 음원과 아티스트의 특성을 반영한다면 음원 또는 아티스트의 특성에 대하여 전략적 측면에서 음원 발매와 순위권 유지에 대한 새로운 관점의 전략을 도출하는데 도움을 줄 수 있을 것으로 판단된다. 두 번째, 음원의 범위를 국내로 한정하여 분석을 수행하였다는 점이다. 디지털 음원의 소비 패턴은 각 국가의 문화적 환경에 따라서 변칙적인 특징을 가질 수 있다. 따라서 이러한 복잡한 유형의 패턴들은 다양한 국가의 음원들을 고려한다면 본 연구에서 제시한 것 이상으로 다양하게 분류될 수 있을 것이다. 마지막으로 각 패턴들의 시계열적 특징에 대한 통계적 검증은 포함되지 않

았다. 시계열을 가지는 데이터들에 대하여 보편적으로 시계열 분석 기법을 적용하며, 이를 통하여 시계열이 가지는 주기, 추세, 계절성 등의 시계열적 특성들을 검증해 볼 수 있다. 따라서 음원의 경우 계절에 의한 영향을 크게 받을 수 있는 콘텐츠로 음원 발매 시점의 계절과 시계열적 특성을 검증하기 위하여 시계열 분석 기법을 적용한다면 시각적으로는 잘 보이지 않는 패턴의 특징을 확인하여 보다 풍부한 시사점을 제시할 수 있을 것으로 기대된다.

## 참고문헌(References)

- Adler, M., "Stardom and talent," *The American economic review*, Vol.75, No.1, (1985), 208~212.
- Berry, M., Linoff, G., *Data mining techniques: For marketing*, (1997), Sales and Marketing Support.
- Bhattacharjee, S., R. D. Gopal, K. Lertwachara, J. R. Marsden, and R. Telang, "The effect of digital sharing technologies on music markets: A survival analysis of albums on ranking charts," *Management Science*, Vol.53, No.9, (2007), 1359~1374.
- Bishop, C. M., *Pattern recognition and machine learning*, (2006), springer.
- Cheom A. -Y., "A Study on the Success Factors of Popular Music Sources," *Master's Thesis*, Sungkyunkwan University Graduate School, 2018.
- Cuturi, M., "Fast global alignment kernels," *In Proceedings of the 28th international conference on machine learning (ICML-11)*, (2011), 929~936.

- Han, E. J., "A Study on the Relationship between Popularity of Advertising Background Music and Musical Elements," *Master's Thesis*, Graduate School of Journalism and Broadcasting, Chung-Ang University, 2004.
- Hyeon, C. -M., "A Study on the Success Factors of Digital Sound Source," *Master's Thesis*, Graduate School, Chung-Ang University, 2014.
- Jo, B. C. and H. C. Sim, "Success Factor Analysis of K-POP and A Study on sustainable Korean Wave - Focus on Smart Media based on Realistic Contents," *The Journal of the Korea Contents Association*, Vol.13, No.5, (2013), 90~102.
- Jo, S. M., "Searching for revitalization of global expansion through social networks (SNS) in the digital music market," *Master's thesis*, Graduate School of Joong-bu University, 2013.
- Jung, E. Y., "A Study on New Ecosystem Model through Case Comparison of Korean, US, and Japanese Digital Music Markets," *Ph.D. Thesis*, Graduate School of Arts, Chung-Ang University, 2016.
- Kim, H. J., "Star Power Analysis of Korean Movie Stars," *Korea Association for Cultural Economics*, Vol.1, No.1, (1998), 165~200.
- Kim, H. -S., "A Study on the Influence and Development Direction of Digital Technology on the Music Industry," *Master's Thesis*, Graduate School of Media and Public Relations, Yonsei University, 2002.
- Lee, S. M., "A Study on Marketing of Popular Music Using SNS," *Master's Thesis*, Graduate School of Arts, Chung-Ang University, 2019.
- Lee, W. G., I. S. Han, and Y. M. Yoon, "Prediction of Record Charts Progress," *Journal of Korean Institute of Information Technology*, Vol.12, No.3, (2014), 121~128.
- Ministry of Culture, Sports and Tourism, *2018 Overseas Korean Wave Survey Report*, Ministry of Culture, Sports and Tourism, 2018.
- Montero, P., Vilar, J. A., "TSclust: An R package for time series clustering," *Journal of Statistical Software*, Vol.62, No.1, (2014), 1~43.
- Moon, H. -C., "A Study on Changes in Market Environment by Convergence of Music Industry and Digital Technology," *Master's Thesis*, Graduate School of Business, Dankook University, 2011.
- Pachet, F., "Hit Song Science. Pitch, harmony, and neural nets: A psychological perspective," *Music and Data Mining*, (2011), 306~314.
- Pachet, F., and P. Roy, "Hit Song Science Is Not Yet a Science," *In ISMIR*, (2008), 355~360.
- Ryu, J. -S., "A Study on the Current Status of the Digital Music Market," *Master's Thesis*, Graduate School of Culture and Arts, Dankook University, 2014.
- Strobl, E. A., and C. Tucker, "The dynamics of chart success in the UK pre-recorded popular music industry," *Journal of Cultural Economics*, Vol.24, No.2, (2000), 113~134.
- Wilks, D. S., *Statistical methods in the atmospheric sciences*, Vol.100, (2011), Academic press.
- Adler, M., "Stardom and talent," *The American economic review*, Vol.75, No.1, (1985), 208~212.
- Berry, M., Linoff, G., *Data mining techniques: For marketing*, (1997), Sales and Marketing Support.



Bhattacharjee, S., R. D. Gopal, K. Lertwachara, J. R. Marsden, and R. Telang, "The effect of digital sharing technologies on music markets: A survival analysis of albums on ranking charts," *Management Science*, Vol.53, No.9, (2007), 1359~1374.

Bishop, C. M., *Pattern recognition and machine learning*, (2006), springer.

Cheom A. -Y., "A Study on the Success Factors of Popular Music Sources," *Master's Thesis*, Sungkyunkwan University Graduate School, 2018.

Cuturi, M., "Fast global alignment kernels," *In Proceedings of the 28th international conference on machine learning (ICML-11)*, (2011), 929~936.

Han, E. J., "A Study on the Relationship between Popularity of Advertising Background Music and Musical Elements," *Master's Thesis*, Graduate School of Journalism and Broadcasting, Chung-Ang University, 2004.

Hyeon, C. -M., "A Study on the Success Factors of Digital Sound Source," *Master's Thesis*, Graduate School, Chung-Ang University, 2014.

Jo, B. C. and H. C. Sim, "Success Factor Analysis of K-POP and A Study on sustainable Korean Wave - Focus on Smart Media based on Realistic Contents," *The Journal of the Korea Contents Association*, Vol.13, No.5, (2013), 90~102.

Jo, S. M., "Searching for revitalization of global expansion through social networks (SNS) in the digital music market," *Master's thesis*, Graduate School of Joong-bu University, 2013.

Jung, E. Y., "A Study on New Ecosystem Model through Case Comparison of Korean, US,

and Japanese Digital Music Markets," *Ph.D. Thesis*, Graduate School of Arts, Chung-Ang University, 2016.

Kim, H. J., "Star Power Analysis of Korean Movie Stars," *Korea Association for Cultural Economics*, Vol.1, No.1, (1998), 165~200.

Kim, H. -S., "A Study on the Influence and Development Direction of Digital Technology on the Music Industry," *Master's Thesis*, Graduate School of Media and Public Relations, Yonsei University, 2002.

Lee, S. M., "A Study on Marketing of Popular Music Using SNS," *Master's Thesis*, Graduate School of Arts, Chung-Ang University, 2019.

Lee, W. G., I. S. Han, and Y. M. Yoon, "Prediction of Record Charts Progress," *Journal of Korean Institute of Information Technology*, Vol.12, No.3, (2014), 121~128.

Ministry of Culture, Sports and Tourism, *2018 Overseas Korean Wave Survey Report*, Ministry of Culture, Sports and Tourism, 2018.

Montero, P., Vilar, J. A., "TSclust: An R package for time series clustering," *Journal of Statistical Software*, Vol.62, No.1, (2014), 1~43.

Moon, H. -C., "A Study on Changes in Market Environment by Convergence of Music Industry and Digital Technology," *Master's Thesis*, Graduate School of Business, Dankook University, 2011.

Pachet, F., "Hit Song Science. Pitch, harmony, and neural nets: A psychological perspective," *Music and Data Mining*, (2011), 306~314.

Pachet, F., and P. Roy, "Hit Song Science Is Not Yet a Science," *In ISMIR*, (2008), 355~360.

Ryu, J. -S., "A Study on the Current Status of the Digital Music Market," *Master's Thesis*, Graduate School of Culture and Arts, Dankook University, 2014.

Strobl, E. A., and C. Tucker, "The dynamics of

chart success in the UK pre-recorded popular music industry," *Journal of Cultural Economics*, Vol.24, No.2, (2000), 113~134.

Wilks, D. S., *Statistical methods in the atmospheric sciences*, Vol.100, (2011), Academic press.

Abstract

## Derivation of Digital Music's Ranking Change Through Time Series Clustering

In-Jin Yoo\* · Do-Hyung Park\*\*

This study focused on digital music, which is the most valuable cultural asset in the modern society and occupies a particularly important position in the flow of the Korean Wave. Digital music was collected based on the “Gaon Chart,” a well-established music chart in Korea. Through this, the changes in the ranking of the music that entered the chart for 73 weeks were collected. Afterwards, patterns with similar characteristics were derived through time series cluster analysis. Then, a descriptive analysis was performed on the notable features of each pattern. The research process suggested by this study is as follows. First, in the data collection process, time series data was collected to check the ranking change of digital music. Subsequently, in the data processing stage, the collected data was matched with the rankings over time, and the music title and artist name were processed. Each analysis is then sequentially performed in two stages consisting of exploratory analysis and explanatory analysis. First, the data collection period was limited to the period before ‘the music bulk buying phenomenon’, a reliability issue related to music ranking in Korea. Specifically, it is 73 weeks starting from December 31, 2017 to January 06, 2018 as the first week, and from May 19, 2019 to May 25, 2019. And the analysis targets were limited to digital music released in Korea. In particular, digital music was collected based on the “Gaon Chart”, a well-known music chart in Korea. Unlike private music charts that are being serviced in Korea, Gaon Charts are charts approved by government agencies and have basic reliability. Therefore, it can be considered that it has more public confidence than the ranking information provided by other services. The contents of the collected data are as follows. Data on the period and ranking, the name of the music, the name of the artist, the name of the album, the Gaon index, the production company, and the distribution company were collected for the music that entered the top 100 on the music chart within the collection period. Through data collection, 7,300 music, which were included in the top 100 on the music chart, were

---

\* Graduate School of Business IT, Kookmin University

\*\* Corresponding author: Do-Hyung Park

School of MIS / Graduate School of Business IT, Kookmin University

77, Jeongneung-ro, Seongbuk-gu, Seoul, Republic of Korea

Tel: +82-2-910-5613, E-mail: dohyungpark@kookmin.ac.kr

identified for a total of 73 weeks. On the other hand, in the case of digital music, since the cases included in the music chart for more than two weeks are frequent, the duplication of music is removed through the pre-processing process. For duplicate music, the number and location of the duplicated music were checked through the duplicate check function, and then deleted to form data for analysis. Through this, a list of 742 unique music for analysis among the 7,300-music data in advance was secured. A total of 742 songs were secured through previous data collection and pre-processing. In addition, a total of 16 patterns were derived through time series cluster analysis on the ranking change. Based on the patterns derived after that, two representative patterns were identified: ‘Steady Seller’ and ‘One-Hit Wonder’. Furthermore, the two patterns were subdivided into five patterns in consideration of the survival period of the music and the music ranking. The important characteristics of each pattern are as follows. First, the artist's superstar effect and bandwagon effect were strong in the one-hit wonder-type pattern. Therefore, when consumers choose a digital music, they are strongly influenced by the superstar effect and the bandwagon effect. Second, through the Steady Seller pattern, we confirmed the music that have been chosen by consumers for a very long time. In addition, we checked the patterns of the most selected music through consumer needs. Contrary to popular belief, the steady seller: mid-term pattern, not the one-hit wonder pattern, received the most choices from consumers. Particularly noteworthy is that the ‘Climbing the Chart’ phenomenon, which is contrary to the existing pattern, was confirmed through the steady-seller pattern. This study focuses on the change in the ranking of music over time, a field that has been relatively alienated centering on digital music. In addition, a new approach to music research was attempted by subdividing the pattern of ranking change rather than predicting the success and ranking of music.

**Key Words** : Music, Digital Music, Rank Change, Time Series Clustering

Received : July 14, 2020 Revised : September 25, 2020 Accepted : September 26, 2020

Corresponding Author : Do-Hyung Park

## 저자 소개



### 유인진

국민대학교 경영정보학부에서 학사 학위를 취득하였으며, 현재 국민대학교 비즈니스 IT 전문대학원에서 CX Lab.에 소속되어 Customer Experience, Business Analytics 트랙으로 박사과정에 재학 중이다. 주요 관심 분야는 Customer Behavior & Analytics이며, 그 외 SME, R&D, Trading area 등에 대한 정량적, 정성적 분석 등을 수행하고 있다.



### 박도형

KAIST 경영대학원에서 MIS 전공으로 석사 / 박사학위를 취득하였다. 현재 국민대학교 경영대학 경영정보학부 / 비즈니스 IT 전문대학원 부교수로 재직 중이며, 고객경험연구실 (CXLab.)을 책임지고 있다 ([www.cxlab.co.kr](http://www.cxlab.co.kr)). 한국 과학 기술 정보 연구원(KISTI)에서 유망아이템 발굴, 기술가치 평가 및 로드맵 수립, 빅데이터 분석 등을 수행하였고, LG 전자에서 통계, 시선/뇌파 분석, 데이터 마이닝을 활용한 소비자 평가 모형 개발을 담당하였고, 스마트폰, 스마트TV, 스마트Car 등에 대한 Technology, Business, Market Insight 기반 컨셉 도출 프로젝트를 다수 수행하였다. 현재 주요 관심분야는 사회심리학 기반의 사용자/소비자의 행동 이론(User/Customer Behavior), 통계 및 인공지능 기법 기반의 사

용자/소비자 애널리틱스 (User/Customer Analytics), 디자인사고 (Design Thinking) 기반의 사용자/소비자 경험 디자인 (Experience Design)이다.