

장바구니 구조와 전환 행동이 소비자 가치에 미치는 영향*

황자성**, 이로***, 김태완****

본 연구는 장바구니 수준의 거래 데이터를 기반으로, 장바구니 구조와 그 변화가 소비자의 소비 패턴에 어떠한 영향을 미치는지를 분석한다. 기존 연구가 주로 단일 구매 결과에 초점을 맞춘 것과 달리, 본 연구는 군집분석을 통해 네 가지 대표적인 장바구니 유형을 도출하고, 개인 수준에서 주도적인 장바구니 스타일과 장바구니 스타일 전환 행동을 구성하여 소비자의 안정적인 소비 성향과 동적 행동 변화를 함께 반영한다. 또한, 서로 다른 소매업체 및 채널 간 선택 범위와 같은 소비자가 이용하는 소비 환경의 다양성을 도입하여 그 조절 효과를 분석한다. 실증 결과는 다음과 같다. 먼저, 다양형 또는 조합형 장바구니를 주로 보이는 소비자는 소비 금액이 높게 나타난다. 또한, 장바구니 스타일 전환 횟수는 세 가지 모델에서 유의한 영향을 보이며, 더 높은 구매 빈도, 더 큰 소비 금액, 그리고 더 짧은 재구매 간격과 연관된다. 이러한 결과는 소비자 가치가 단순한 소비 수준뿐 아니라, 서로 다른 장바구니 구조 간 전환 능력과도 밀접하게 관련되어 있음을 시사한다. 본 연구는 “장바구니가 어떻게 구성되는가”와 “장바구니가 어떻게 변화하는가”라는 두 가지 관점에서 접근하여, 소비자 행동을 이해하기 위한 새로운 분석 틀을 제시한다. 동시에, 옴니채널 환경에서 소매업체의 고객 관리 전략에 대해 구체적인 시사점을 제공한다.

주제어 : 장바구니 기반 소비행동, 군집 분석, 고객가치

I. 서론

전자상거래와 데이터 기록 기술의 발전에 따라, 연구자들은 이전에는 불가능했던 수준의 세밀한 해상도로 소비자의 구매 행동을 관찰할 수 있게 되었다. 대규모 거래 데이터가 연구에 활용될 수 있는 상황은 소비자 행동 연구를 설문조사와 집계 지표에 의존하던 거시적 관점에서, 실제 거래에 기반한 미시적 분석으로 점진적으로 전환시키고 있다(Li & Kannan, 2014). 이 가운데, ‘장바구니(basket)’를 분석 단위로 하는 데이터 구조는 특히 중요한 연구 대상이다. 각각의 장바구니는 소비자가 한 번의 구매 행동에서 선택한 상품 조합

을 의미하며, 특정 시점에서의 선호 구조와 의사 결정 논리를 반영한다(Agrawal et al., 1993). 거래 빈도나 금액만을 지표로 사용하는 기존 연구와 비교할 때, 장바구니 수준의 분석은 소비자가 다품목을 어떻게 조합하는지, 예산을 어떻게 배분하는지, 그리고 구매 품목들이 어떻게 연결되는지를 보다 복합적으로 드러낼 수 있다(Brijs et al., 1999).

기존의 장바구니 관련 문헌을 검토해보면, 여전히 세 가지 뚜렷한 연구 공백을 확인할 수 있다. 첫째, 소비자 분류 방식이 정태적이라는 점이다. 전통적인 시장 세분화는 주로 성별, 연령, 소득과 같은 인구통계학적 변수에 근거해 왔으며(Kotler & Keller, 2016), 이러한 접근은 거시적 특성을 파

* 이 논문은 2025년 한국유통학회와 텐큐브의 학술데이터지원사업 지원을 받아 수행된 연구임.

** 건국대학교 경영학과 박사과정 수료(huangzixing711@konkuk.ac.kr), 제1저자

*** 건국대학교 경영학과 박사과정(lilu1999@konkuk.ac.kr), 공동저자

**** 건국대학교 경영학과 교수(tkim21@konkuk.ac.kr), 교신저자

약하는 데에는 유용하지만 소비자의 행동 방식이나 의사결정 리듬에서 나타나는 동태적 차이를 포착하는 데에는 한계가 있다. 즉, 소비자가 단지 ‘누구인가’를 보는 것이 아니라, 소비자가 ‘어떻게 소비하는가’, ‘언제 소비하는가’, ‘왜 소비하는가’라는 차원에서 다양한 동태적 구매 행동 양상을 보인다.

둘째, ‘장바구니’에 관한 연구는 여전히 단일 거래 분석에 국한되어 있다. 초기의 ‘시장 장바구니 분석’은 주로 상품 수준에서 동시에 구매하는 규칙에 초점을 맞추어(Agrawal et al., 1993), 어떤 상품들을 함께 구매하는지를 밝혀내는 데 기여했으나, 동일한 소비자가 서로 다른 거래 시점에서 보이는 행동의 변화와 시간에 따라 진화하는 현상은 충분히 고려하지 못했다.

셋째, 최근 등장한 장바구니 시퀀스 연구는 시간 차원을 도입하여 소비자 행동의 동태성에 주목하기 시작했으나, 이러한 연구 목표는 주로 추천 시스템 맥락에 집중되어 있다(Rendle et al., 2010). 이들 연구는 대체로 ‘다음에 무엇을 구매할 것인가’를 예측하는 데 초점을 두고 단기 예측 정확도를 강조하며, 장바구니 사이의 구조적 변화가 더 긴 시간 범위에서 안정적인 행동 결과로 누적되고 분석되는지에 대해서는 논의된 바가 없다. 실제로 장바구니 간의 변화는 선호의 진화 과정을 반영할 뿐만 아니라, 구매 근접성(recency), 거래 빈도, 지출액 같은 소비자의 장기적 행동 결과에도 영향을 미칠 수 있다.

본 연구는 ‘거래 수준 장바구니’를 분석 단위로 연구한다. 동일한 소비자가 동일한 시점에 동일한 소매점에서 한 차례 거래를 완료한 경우, 구매한 상품의 수나 종류와 관계없이 이를 하나의 완전한 장바구니로 정의한다. 이러한 정의는 ‘시간 일관성’과 ‘소매점 일관성’을 핵심 기준으로 하여, 단일 실제 구매 상황에서 소비자가 수행한 의사

결정 행동을 최대한 충실하게 반영할 수 있도록 한다. ‘동일 시간-동일 소매점’을 기준으로 한 이 정의를 통해, 본 연구는 미시적 수준에서 단일 소비 행위의 내부 구조(예: 상품 수, 범주 다양성, 지출 구성)를 파악할 수 있을 뿐만 아니라, 종단적 관점에서 동일 소비자가 서로 다른 거래 시점에서 형성하는 장바구니의 변화를 연속적으로 추적할 수 있다. 이러한 설계는 연구가 단일 의사결정의 특성을 정태적으로 분석하는 데 그치지 않고, 개인의 소비 선호가 시간에 따라 어떻게 진화하는지, 그리고 그 행동이 얼마나 안정적으로 유지되는지를 동태적으로 식별할 수 있게 한다.

따라서 본 연구는 단일 장바구니를 기반으로 하되, 구조적 특성과 시간적 진화를 함께 고려하는 3단계 분석 프레임워크를 활용한다. 제1단계에서는 단일 거래를 출발점으로 하여 장바구니의 구성 특성에 대해 군집을 분석을 하여, 여러 상황에서 나타나는 소비자의 전형적인 쇼핑 스타일을 식별한다. 제2단계에서는 이러한 쇼핑 스타일을 시간 순서에 따라 연결하여, 소비자 쇼핑 행동의 변화 궤적과 진화 논리를 설명한다. 제3단계에서는 서로 다른 쇼핑 경로가 장기 행동 결과에 어떠한 차이를 가져오는지를 분석하고, 이 과정에서 소매업체 선택의 집중도가 이러한 관계를 어떻게 조절하는지를 함께 검토한다.

이 3단계 프레임워크는 단일 거래나 정태적 분류에 국한된 관점을 넘어, 소비자의 쇼핑 행동을 연속적이며 진화 가능한 과정으로 이해할 수 있게 한다. 장바구니의 구조적 특성과 시간 차원에서의 변화를 연결하여, 반복 구매 과정에서 소비자가 어떻게 점차 자신만의 안정적인 구매행동 리듬과 의사결정 논리를 형성해 가는지를 설명할 수 있다. 이러한 관점은 기존 문헌이 지니는 정태적 세분화, 단일 장바구니 분석, 그리고 장기 행동 결과와의 연계 부족이라는 한계를 보완하고, 기업

이 거래 데이터를 기반으로 보다 정교하고 목적 지향적인 소비자 인사이트를 도출하고 마케팅 전략을 고도화할 수 있게 한다.

II. 이론적 배경

1. 사용자 행동 모델링 연구: 정태적 이미지에서 동태적 궤적으로

기존 연구는 주로 인구통계학적 특성과 심리적 변수를 통해 개인 간 차이를 설명해 왔다. Kotler and Keller(2016)가 제시한 시장 세분화 모형은 이러한 정태적 분류 접근의 전형적인 예로, 연구자들은 연령, 성별, 소득, 라이프스타일 등의 변수를 활용하여 시장을 비교적 동질적인 잠재 고객 집단으로 구분하였다. Engel-Kollat-Blackwell 모형은 한 걸음 더 나아가 심리적 과정에 기반해 소비자의 의사결정 논리를 설명하고자 하였다. 이러한 모형들은 전통적 소매 환경에서는 이론적 의미가 있지만, 일반적으로 소비자의 특성이 비교적 안정적이라고는 가정하고 있다.

그러나 디지털화와 모바일화가 진전된 소비 환경에서는 이러한 가정이 점차 유효성을 잃고 있다. 소비자 구매 행동은 빈도가 매우 높은 편이고 동태적이며, 다양한 플랫폼을 넘나드는 형태로 나타나고 있으며, 기존의 정태적 소비자 분석으로는 이러한 빠른 변화 양상을 충분히 포착하기 어렵다(Wedel & Kamakura, 2000). 이러한 배경에서 연구의 초점은 점차 ‘누가 소비하는가’에서 ‘어떻게 소비하는가’로 이동하였고, 사용자 행동 모형 역시 정태적 특성 중심에서 행동 중심 접근으로 전환되기 시작하였다.

RFM(recency-frequency-monetary) 모형은 최근 구매 시점, 구매 빈도, 소비 금액을 통해 사용자의

활동성과 잠재 가치를 측정하며, 복잡한 심리 측정 없이도 핵심 고객을 효율적으로 식별할 수 있게 한다(Tsai & Chiu, 2004). RFM 모형의 도입은 사용자 행동을 정태적 특성에서 벗어나 측정 가능한 행동 지표로 전환시켰다는 점에서 의의가 있으며, 이후 동적 군집화 및 행동 기반 세분화 연구의 기초를 제공하였다.

후속 연구들은 데이터 마이닝과 머신러닝 방법의 발전에 따라, RFM 프레임워크에 군집분석과 동적 세분화 기법을 도입하여 사용자 행동을 지속적으로 추적하고 계층적으로 관리하고자 하였다(Abbasimehr & Bahrini, 2022; Wang et al., 2024). 이는 소비자가 더 이상 정태적인 집단 라벨이 아니라, 시간의 흐름 속에서 지속적으로 변화하는 행동 주제로 이해되기 시작했음을 의미한다. 이러한 맥락에서 사용자 행동 데이터가 축적됨에 따라, 점차 더 많은 연구자들이 ‘행동 시퀀스(behavioral sequence)’라는 새로운 분석 차원에 주목하게 되었다.

기존 연구에 따르면 소비자의 구매 행동은 시간 의존성과 상태 전이 특성을 지니는 경우가 많으며, 단순히 정태적 변수에 기반한 군집 모형으로는 이러한 동적 진화를 충분히 설명하기 어렵다(Hidasi et al., 2015). 이에 따라 순환 신경망(RNN), 장단기 기억 네트워크(LSTM), Transformer와 같은 딥러닝 모형이 전자상거래 및 소매 환경에 폭넓게 도입되었으며, 이들 모형은 행동 시퀀스 내에서 장기적·단기적 선호 변화 모두를 포착하는데 활용되고 있다(Tan, Xu, & Liu, 2016). 이러한 접근은 연속적인 구매, 탐색, 클릭 시퀀스를 학습함으로써 소비자 선호의 미세한 변화를 식별할 수 있게 하며, 그 결과 소비 추천이나 다음 구매 예측에서 더 높은 정확성과 유연성을 보여준다. 비록 이들 연구가 주로 추천 시스템 맥락에 집중되어 있기는 하나, 그 기본적인 접근방식은 소비

자가 여러 차례의 구매 의사결정을 거치며 어떻게 행동을 변화시키는지를 이해하는 데에도 중요한 시사점을 제공한다.

2. 거래 장바구니 분석: 구조적 소비 스타일의 식별

‘장바구니’ 개념은 장바구니 분석(MBA: market basket analysis)에서 기원하였으며, 그 핵심 목적은 거래 데이터를 활용하여 상품 간 동시에 구매하는 규칙을 발굴하는 데 있다(Agrawal et al., 1993). 예를 들어, 슈퍼마켓 데이터에서는 ‘우유를 구매한 고객이 빵도 함께 구매하는 경향이 있다’는 패턴이 자주 관찰되며, 기업은 이를 바탕으로 상품 진열 최적화나 묶음 판촉 전략을 설계할 수 있다. 이러한 연구들은 소매 데이터 분석의 기초를 마련하였으나, 분석 단위가 ‘상품-상품’ 간 동시에 구매한 관계에 집중되어 있어 정태적인 구매 연관성만을 반영할 뿐, 소비자가 한 번의 완결된 거래에서 어떠한 의사결정 논리와 구매 스타일을 보이는지는 충분히 설명하기 어렵다.

본 연구에서 정의한 거래 수준 장바구니를 기반으로, 후속 연구들은 분석 단위를 단일 거래의 구조적 특성으로 확장해 왔다. 이에 따라 연구의 초점 역시 ‘상품이 어떻게 함께 구매되는가’에서 ‘소비자가 한 번의 쇼핑에서 자원을 어떻게 배분하고, 구매 선택을 어떻게 구성하며, 선호를 어떻게 표현하는가’로 이동하였다. 종합적으로 볼 때, 선행연구들은 본 연구에서 제시하는 3C 프레임워크, 즉 구성(composition), 복잡성(complexity), 상황(context)에 대해 이론적·실증적 기반을 제공한다. 이 세 가지 차원에서 장바구니의 구조적 특성을 규정함으로써, 거래 수준에서 서로 다른 소비 스타일 유형을 파악할 수 있으며, 나아가 이러한 구조적 차이가 소비자의 장기적 가치와 어떻게

연관되는지를 탐구할 수 있는 토대가 마련된다.

(1) 거래 구성은 단일 거래에서의 ‘수량과 가격’ 구조를 강조한다. 즉, 소비자가 한 번의 구매에서 얼마나 많은 상품을 구매했는지, 얼마를 지출했는지, 그리고 그 지출 규모가 어떠한 소비 강도를 나타내는지 의미한다. 다시 말해, 이는 ‘이 장바구니가 무엇으로 구성되어 있는가’라는 질문에 답한다. 기존 연구에 따르면, 단일 거래에서의 구매 수량과 지출 수준은 소비자의 쇼핑 목적과 과업 특성을 드러내는 중요한 단서가 된다. 예를 들어, Boztug and Reutterer(2008)는 장바구니의 금액 구조와 구성 특성을 기반으로 군집분석을 수행하여 ‘고가 집중형’, ‘저가 다품목형’, ‘보충형’, ‘비축형’ 등과 같은 전형적인 소비 패턴으로 구분하였으며, 이를 통해 소비자들이 단일 거래에서 유의미하게 상이한 행동을 보인다고 하였다. 따라서 구조적 군집분석에서 상품 수량과 지출 금액과 같은 변수를 모형에 포함하는 것은, 서로 다른 쇼핑 스타일과 과업 강도 간의 잠재적 차이를 구분하는 데 도움이 된다(Bell et al., 2011).

(2) 거래 복잡성은 소비자가 ‘무엇을 구매하는가’를 이해한 이후, 연구자들은 나아가 ‘얼마나 복합적으로 구매하는가’에 주목하기 시작하였다. 복잡성(complexity) 차원은 단일 거래 내부의 다양성과 집중도를 측정하며, 이는 소비자가 구매할 때 지출을 분산하여 여러 선택지를 탐색하는지, 혹은 자원을 집중하여 특정 대상에 목적 지향적으로 구매하는지를 반영한다. Guidotti et al.(2015)은 처음으로 ‘장바구니 엔트로피’라는 개념을 제시하고, Shannon 엔트로피를 활용하여 장바구니 내부의 불확실성과 다양성을 정량화하였다. 엔트로피 값이 높을수록 지출이 여러 상품에 비교적 고르게 분포되어 소비자의 탐색 성향이 강함을 의미하며, 반대로 엔트로피 값이 낮을수록 구매가 특정 대상에 집중되어 목표 지향적이고 계획적인

특성이 두드러짐을 나타낸다. 선행연구에 따르면, 엔트로피 값이 높은 소비자는 장기적인 소비 과정에서 더 높은 활동성과 다양화된 구매 성향을 보이는 경향이 있으며(Guidotti et al., 2015), 이러한 탐색적 행동은 잠재적으로 높은 가치의 고객을 예측하는 초기 신호로 해석되어 왔다(McAlister & Pessemier, 1982). 다시 말해, 장바구니의 복잡성은 소비자의 쇼핑 스타일이 계획적 구매인지 탐색적 시도인지를 드러낼 뿐만 아니라, 향후 소비 잠재력과도 연관성을 지닌다. 따라서 거래 수준의 구조 분석에서 복잡성(complexity)을 모형에 포함시키는 것은, 소비자가 단일 구매 상황에서 주의와 자원을 어떻게 배분하는지를 이해하는 데 기여하며, 이를 통해 소비자의 행동 특성과 가치 추이를 보다 입체적으로 포착할 수 있게 한다.

(3) 거래 상황 차원은 쇼핑 행동이 발생하는 시간적·공간적·채널적 환경에 주목한다. 쇼핑은 고립된 사건이 아니라, 소비자의 일상생활 속에 내재된 동적인 과정이다. 선행연구에서, 구매가 이루어지는 시간대(예: 아침 또는 야간), 주말 여부(평일 또는 주말), 그리고 쇼핑 채널(온라인 또는 오프라인)은 모두 거래 구조에 유의미한 영향을 미치는 것으로 나타났다(Bell et al., 2011; Hui et al., 2013). 예를 들어, 평일 오전의 쇼핑은 대체로 보충 구매나 계획적인 구매가 중심이 되는 반면, 주말 야간의 쇼핑은 보다 쾌락적이거나 즉흥적인 탐색 성향을 띠는 경우가 많다. 이러한 차이는 쇼핑의 리듬이 생활의 리듬과 밀접하게 연관되어 있음을 시사한다. 소비자는 서로 다른 시간대에서 상이한 정서 상태, 시간 압박, 사회적 환경의 영향을 받으며 구매 결정을 내린다. 이처럼 파편화되고 다중 맥락으로 이루어진 소비 방식이 확산됨에 따라, 소비자의 의사결정 논리는 더 이상 단일 거래에 국한하여 이해될 수 없으며, 연속적인 행동 궤적 속에서 분석될 필요가 있다. Lemon and

Verhoef(2016)가 제시한 ‘고객 여정’ 이론은 이러한 관점에 이론적 근거를 제공한다. 이들은 각 거래를 그것이 발생한 구체적인 시간, 장소, 채널과 같은 맥락 속에서 분석해야 소비자의 심리적 동기와 의사결정 메커니즘을 정확히 설명할 수 있다고 주장한다. 따라서 맥락(context) 차원을 구조 분석에 포함시키는 것은 ‘언제 구매하는가’라는 외형적 규칙을 밝히는 데 그치지 않고, ‘왜 그러한 방식으로 구매하는가’라는 내적 논리를 이해하는 데에도 도움을 준다.

종합하면, 이 세 가지 차원은 함께 거래 수준 장바구니의 구조화된 이미지를 구성하며, 이를 통해 서로 다른 소비 스타일과 행동 논리를 하나의 통합된 틀에서 분석할 수 있게 한다.

기존 연구들은 주로 단일 거래의 구조를 중심으로 소비 패턴을 설명해 왔으며, 이로 인해 반복 구매 과정에서 나타나는 소비자의 동적 행동 변화를 충분히 반영하지 못하고 있다. 소비자가 여러 차례의 거래를 거치면서 보이는 행동의 진화를 충분히 고려하지 못하고 있다. 반복 구매 과정에서 소비자는 광범위한 탐색 단계에서 출발하여 선호의 고착화, 나아가 안정적인 반복 구매로 이행하는 변화를 경험할 수 있다. McAlister and Pessemier(1982)의 초기 실증 연구는 신규 고객이 초기에는 다양한 브랜드와 제품을 시도하지만, 경험의 축적과 신뢰의 형성에 따라 선택이 점차 수렴한다는 점을 보여주었다. 이러한 ‘탐색-고착-안정’의 궤적은 소비자 학습과 브랜드 관계 심화의 핵심적 표현으로 해석된다. 그럼에도 불구하고, 대다수 연구는 이러한 행동 양상을 기술하는 데 그치고 있으며, ‘장바구니 내부 구조’와 ‘장바구니 간 변화’를 결합하여 분석하는 체계적 접근은 여전히 부족하다. Reinartz and Kumar(2003)는 고객의 장기 가치를 제대로 이해하기 위해서는 개별 거래를 고립된 사건이 아니라, 연속적인 행동 과

정으로 바라보아야 한다고 지적한다. 동일한 소비자가 여러 장바구니를 거치며 보이는 구조적 변화를 추적할 수 있을 때에야 비로소, 소비 습관이 반복된 선택을 통해 어떻게 점차 고착화되는지, 그리고 그러한 변화가 어떻게 장기적인 고객 가치로 전환되는지를 밝힐 수 있다.

따라서 향후 연구의 방향은 단일 거래의 구조 분석과 다중 거래에 걸친 동적 진화를 결합하는 종합적 모형을 구축하는 데 있다. 이러한 접근은 소비자의 장기 행동을 보다 정교하게 설명할 수 있을 뿐만 아니라, 반복 구매 과정에서 형성되는 소비 패턴을 이해하는 데 기여한다. 이와 같은 구조화된 이해는 행동 분석과 가치 관리 사이를 연결하는 중요한 이론적·실무적 가교가 될 것이다.

3. 장바구니 시퀀스와 사용자 가치 연구: 행동 예측에서 행동 결과 모형화로

단일 장바구니 분석의 정태적 한계를 극복하기 위해, 최근 연구의 초점이 점차 ‘장바구니 시퀀스’의 동적 모형화로 이동되고 있다. 이 연구 흐름의 핵심 목적은 소비자가 여러 차례의 거래를 수행하는 과정에서 장바구니 간에 형성되는 구조적 연계를 이해하는 데 있으며, 특히 ‘다음 장바구니 예측’에 주목한다. 전통적인 연구에서 여러 상품을 동시에 구매한 소비자를 분석할 때 ‘어떤 상품이 함께 구매되는가’에 머무르는 것과 달리, 이들 연구는 ‘소비자가 다음에 무엇을 구매할 가능성이 있는가’와 ‘그러한 구매 패턴이 시간의 흐름에 따라 어떻게 변화하는가’를 보다 직접적으로 탐구한다.

Rendle et al.(2010)은 FPMC(factorizing personalized markov chains) 모델을 통해 개인의 선호와 마르코프 체인을 결합함으로써, 개인 간 이질성과 시계열적 전이를 동시에 포착할 수 있는 틀을 최

초로 제시하였고, 이를 통해 서로 다른 장바구니 간의 전이 규칙을 규명하였다. 이 접근은 이후 시퀀스 모형화 연구의 기초를 마련하였다. Yu et al. (2016)이 제안한 DREAM(dynamic recurrent baskets model) 모델은 순환신경망(RNN) 구조를 도입하여 장바구니 시퀀스에 존재하는 고차 시계열 의존성을 분석 하였다. 딥러닝 기법의 발전과 함께, 연구자들은 행동 시퀀스 내의 ‘핵심 시점’과 ‘유 의미한 사건’에 대한 분석에도 관심을 기울이기 시작하였다. 예를 들어, Bai et al.(2018)은 어텐션(attention) 메커니즘을 도입하여 핵심 거래와 중요한 상품 속성에 대한 가중치를 강화함으로써, 모델이 ‘다음 장바구니가 무엇인가’를 예측하는데 그치지 않고 ‘왜 그 장바구니가 선택되는가’를 설명하였다. 한편, Wang et al.(2015)의 계층적 표현 모델(HRM)과 Zeng et al.(2024)의 계층적 표현 학습 프레임워크는 ‘이중 학습’ 접근을 통해 각각 ‘장바구니 내부 구조’와 ‘장바구니 사이의 시간적 진화(장바구니 간)’를 별도로 모형화한 후 이를 통합하여, 시퀀스 예측의 정확도와 안정성을 향상 시켰다.

비록 앞서 언급한 연구들이 모델 면에서는 지속적으로 발전해 왔지만, 그 연구 목표는 여전히 단기 성과 예측에 집중되어 있다. 대부분의 모형은 클릭률, 구매 적중률, 혹은 다음 거래에 대한 예측 정확도를 핵심 평가 지표로 삼아, 주어진 과거 시퀀스 조건하에서 “다음에 무엇이 발생할 것인가”를 보다 정밀하게 예측하는 데 초점을 맞춘다. Jannach and Adomavicius(2016)는 이러한 단기적 매칭 성과 중심의 연구 패러다임이 추천 효율성 측면에서는 가치가 있지만, 소비자의 다차원적 행동 결과를 이해하는 데에는 분명한 한계가 있다고 지적하였다. 구체적으로, 기존의 장바구니 시퀀스 연구는 대체로 단기간의 시간 척도에서 소비자의 선택 전이를 효과적으로 묘사할 수는 있

으나, 이러한 시퀀스 특성이 더 긴 시간 창 내에서 누적되어 안정적인 다차원 행동 결과로 이어지는지에 대해서는 충분히 분석하지 못하고 있다. 예를 들어, 시퀀스 모형은 다음 구매 내용에 대해서는 효과적인 예측을 수행할 수 있지만, 장기적인 시퀀스 속에서 소비자가 높은 구매 최근성을 지속적으로 유지하는지, 더 빈번한 거래가 발생하는지, 더 높은 전체 지출 수준과 같은 장기 행동 결과를 점진적으로 형성하는지에 대해서는 명확한 답을 제공하지 못하는 경우가 많다.

그러나 기존의 소비자 행동 및 관계 마케팅 연구에 따르면, 이러한 장기적·다차원적 행동 결과는 소비자의 지속적 상호작용 상태를 설명하는데 핵심적인 의미를 지닌다. 먼저, 구매 최근성은 소비자와 소매업체 또는 플랫폼 간 상호작용의 시간적 거리를 반영하며, 관계가 여전히 ‘지속 상태’에 있는지를 판단하는 중요한 지표로 활용된다. 관련 연구에 따르면, 최근 거래 시점으로부터의 시간이 짧을수록 소비자는 심리적·행동적 차원에서 브랜드와의 높은 연관성을 유지하고 있으며, 향후 재구매가 발생할 확률 또한 높아진다(Fader et al., 2005; Gupta et al., 2006). 최근성은 단순한 시간 간격을 의미하는 것이 아니라, 소비자가 해당 소매업체나 플랫폼을 여전히 자신의 ‘일상적 소비 선택 집합’에 포함하고 있는지를 보여주는 직접적인 지표이다. 따라서 높은 최근성은 관계가 아직 단절되지 않았고 행동적 관성이 계속 유지된다고 해석할 수 있다(Verhoef, 2003).

다음으로, 빈번한 구매 행동은 행동 차원에서의 지속적인 상호작용 강도를 나타낸다. 거래 빈도가 높다는 것은 소비자가 더 짧은 시간 안에 반복적인 구매를 수행하고 있음을 의미하며, 이는 해당 소매업체나 플랫폼이 이미 소비자의 일상적 소비 리듬에 포함되어 있음을 보여준다(Bolton et al., 2004). 기존 연구에서 구매 빈도가 고객의 행동의

지속성(behavioral stickiness)과 관계의 안정성을 측정하는 핵심 지표로 널리 활용되어 왔다(Rust et al., 2004).

세 번째로, 높은 지출 수준은 일반적으로 고객 관계의 심화와 신뢰 축적의 결과로 해석된다. 소비자와 소매업체 간 관계가 지속적으로 발전함에 따라, 단일 거래 금액과 전체 소비 규모는 점차 증가하는 경향을 보이는데, 이러한 변화는 관계 안정화 과정에서 형성되는 심리적 동일시와 지각된 위험 감소에 기인하는 경우가 많다(Reinartz & Kumar, 2003).

마지막으로, 소매업체 선택의 집중도는 소비자의 관계 구조를 나타내는 핵심 행동 특성으로, 소비자가 어떠한 방식으로 소매업체를 선택하고 있는지를 보여준다. McAlister and Pessemier(1982)는 관계 초기 단계에서 소비자가 강한 탐색 성향을 보이지만, 경험의 축적과 신뢰 형성을 거치면서 선택 범위가 점진적으로 수렴한다고 하였다. 이후 연구들 역시 이러한 집중화 경향을 충성도 형성과 관계 심화의 중요한 징후로 간주해 왔다(Oliver, 1999). 소비자가 다양한 상황에서 동일한 브랜드나 플랫폼을 반복적으로 선택할 경우, 그 행동은 정서적 친밀감을 나타낼 뿐만 아니라, 기업 입장에서는 보다 안정적인 수익 흐름에 대한 기대를 가능하게 한다(Reicheld & Sasser, 2003). 소매업체 선택의 집중도는 소비자의 관계 구조를 반영함으로써, 행동의 일관성과 안정성 수준을 규정하고, 이에 따라 동일한 장바구니 시퀀스 특성이라 하더라도 장기 행동 결과에 미치는 영향의 크기와 방향을 달라지게 할 수 있다. 따라서 소매업체 선택의 집중도는 장바구니 시퀀스 특성과 장기 행동 결과 간의 관계를 강화하거나 약화시키는 조절 변수로 작용할 수 있다.

추가로 구분할 필요가 있는 점은, 본 연구가 다루는 장기 행동 결과가 지표의 형태 측면에서는

RFM 프레임워크와 표면적인 유사성을 지니고 있으나, 두 접근이 전제하는 연구 맥락과 분석 대상에는 차이가 있다. 전통적인 RFM 모델은 일반적으로 단일 소매업체 또는 단일 브랜드를 분석 배경으로 설정하며, 소비자와 특정 상점 간의 관계가 이미 형성되어 있다는 가정을 핵심 전제로 한다. 이에 따라 연구의 초점은 해당 관계 내에서 고객의 활동 수준과 경제적 기여도를 파악하는데 있다(Fader et al., 2005; Verhoef, 2003). 이러한 맥락에서 RFM는 모두 ‘동일한 하나의 점포를 기준으로 한’ 행동 결과로 정의되며, 특정 상점과 고객 간 관계의 강도를 측정하는 지표로 활용된다.

그러나 실제 소매 환경에서 소비자는 다수의 대체 가능한 소매업체와 채널에 동시에 노출되어 있으며, 초기 단계의 소비 행동이 반드시 특정 상점 하나를 중심으로 전개되는 것은 아니다. Reinartz and Kumar(2003)가 지적하였듯이, 관계가 형성되기 이전의 단계에서 소비자는 일반적으로 탐색 기간을 거치며, 이 시기에는 구매 행동이 여러 브랜드나 소매업체에 걸쳐 분산되어 나타나고, 아직 안정적인 관계 지향성은 보이지 않는다. 이러한 단계에서는 단일 상점 관점에서 산출된 RFM 지표가 소비자의 실제 행동 상태를 충분히 반영하기 어렵다.

이에 따라 이후 연구에서는 분석 시각을 ‘단일 상점-고객 관계’에서 ‘소매 시스템-고객 행동 궤적’으로 확장할 필요성을 점차 강조해 왔다(Lemon & Verhoef, 2016). 이와 같은 관점에서는 RFM이 더 이상 특정 상점에 한정된 관계 결과가 아니라, 소비자가 전체 소매 시스템 내에서 보이는 행동 결과로 해석되며, 그 형성 과정은 일련의 구체적인 거래 의사결정이 시간에 따라 누적된 결과로 이해된다.

본 연구는 ‘거래 계층 장바구니 시퀀스’라는 관점을 통해 소비자가 여러 차례의 거래를 거치며

보이는 구조적 특성의 진화에 주목하고, 이러한 구조 변화가 장기적으로 지출 수준, 구매 빈도, 그리고 관계 집중도의 추이에 어떠한 영향을 미치는지를 분석하고자 한다. 이를 통해 단기적인 ‘다음 장바구니 예측’에서 벗어나, 장기적인 ‘행동 궤적과 가치 생성’에 집중하게 된다. 본 연구의 목적은 소비자가 연속적인 장바구니를 통해 어떻게 자신만의 소비 패턴과 관계 가치를 점진적으로 축적해 나가는지를 이해하는 것이다. 상품 조합의 다양화, 품목 이동, 지출 구조의 변화와 같은 장바구니 간 동적 전환 특성은 선호의 진화를 반영할 뿐만 아니라, 기업에게 새로운 시장 세분화 관점을 제공한다. 장바구니 수준에서 나타나는 행동 유형과 그 변화 경로를 식별함으로써, 기업은 관계가 심화되고 있는 고객과 이탈 또는 전환 국면에 있는 고객을 구분할 수 있으며, 이를 바탕으로 보다 장기 가치 지향적인 고객 관리 전략을 정교하게 설계할 수 있다.

III. 연구 방법

1. 데이터와 연구 설계

1.1 데이터

본 연구에서 사용한 데이터는 한국의 데이터 분석 기업 텐큐브(Tenqube)가 제공하였다. 데이터는 상품 수준의 거래 기록을 최소 관측 단위로 하며, 각 관측치는 특정 소비자가 특정 소매업체에서 구매한 개별 상품이다. 해당 데이터는 거래, 소비자, 상품, 시간 등 다양한 차원의 정보를 포함하고 있어 소매 환경에서의 실제 소비 과정을 비교적 정밀하게 반영한다. 거래 및 소비자 수준에서 각 기록은 소비자 식별자와 주문 번호를 포함한다.

상품 수준에서는 상품명, 상품 단가, 구매 수량이 기록되어 있다. 모든 가격은 원화 기준이다. 시간 차원에서는 거래 발생의 구체적인 날짜가 포함되며, 표본 기간은 2022년 1월부터 2024년 12월까지이다. 이 기간은 COVID-19의 충격 이후 비교적 안정적인 단계에 해당하며, 팬데믹의 영향이 가장 두드러졌던 시기를 포함하지 않는다. 따라서 외부효과로 인한 비정상적인 소비 변동이 연구 결과에 미치는 영향은 작다. 채널 차원에서는 비식별화된 소매업체 코드가 제공되며, 이를 통해 채널 유형을 구분할 수 있다. 구체적으로 A-L은 온라인 소매 플랫폼을 의미하고, X-Z는 오프라인 소매 채널을 의미한다.

1.2 연구 설계

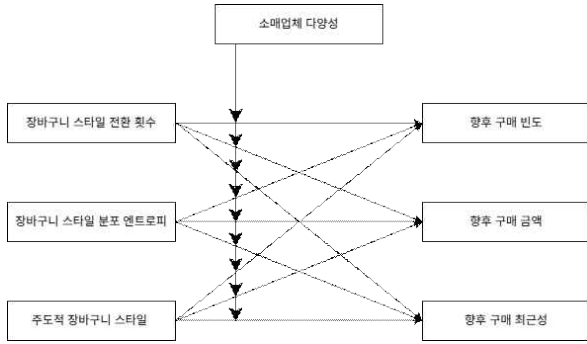
실증 분석에 앞서, 본 연구는 원시 상품 수준의 거래 데이터를 정제하고 “각 상품 기록”을 “하나의 완전한 구매 행위”, 즉 장바구니 단위로 재구성한다. 원시 데이터는 상품 단위로 기록되어 있어, 하나의 관측치는 소비자가 특정 소매업체에서 구매한 단일 상품에만 대응한다. 따라서 소비자의 한 번의 구매는 일반적으로 여러 개의 기록으로 분할되어 나타난다.

실제 소비 행위를 보다 정확하게 반영하기 위해, 본 연구는 “소비자 번호+소매업체 번호+구매 시점(buyat)”을 기준으로 데이터를 연결한다. 동일한 소비자, 동일한 소매업체, 그리고 구매 시점이 완전히 일치하는 여러 기록은 동일한 구매로 간주하며, 하나의 장바구니로 통합한다. 이러한 처리 방식은 데이터에 포함된 시간 정보를 전적으로 활용하여 구매 행위를 정의하며, 별도의 인위적인 시간 구간을 설정하지 않는다. 따라서 주관적인 기준 설정에서 발생할 수 있는 편의를 최소화할 수 있다. 이 과정을 통해 원시의 상품 수준

데이터는 장바구니 단위의 데이터 구조로 변환된다. 각 장바구니는 특정 시점에 특정 소매업체에서 이루어진 하나의 완전한 구매 행위를 의미하며, 해당 구매에 포함된 모든 상품 정보를 함께 포함한다.

장바구니 구성이 완료된 이후, 본 연구는 이에 대한 특성화를 추가로 수행하고 이를 바탕으로 다양한 쇼핑 행태 유형을 식별한다. 구체적으로, 장바구니 수준에서 상품 수, 가격 수준, 상품 구성 구조, 구매 시간 및 채널 등 다차원적 특성을 통합하여 각 장바구니를 정량적으로 묘사한다. 이를 기반으로 군집분석을 적용하여 장바구니를 여러 유형으로 분류하고, 대표적인 특성을 지닌 장바구니 스타일을 도출한다. 이러한 스타일은 구매 구조와 소비 상황에서 뚜렷한 차이를 보인다. 예를 들어, 일부는 소수의 고가 상품을 집중적으로 구매하는 형태를 보이며, 일부는 다양한 저가 상품을 조합하는 소비 형태를 나타낸다. 또한 특정 시간대나 특정 채널 환경에서 주로 발생하는 유형도 존재한다. 본 연구는 이러한 장바구니 수준의 스타일 식별 결과를 개인 수준으로 집계하여 소비자 행동 패턴을 설명하는 핵심 변수를 구성한다. 주도적 스타일, 스타일 전환 빈도, 스타일 분포 엔트로피 등이다. 이를 통해 “단일 구매 행위의 특성”을 “개인 수준의 행동 패턴 지표”로 전환하며, 이후 소비자 행동의 지속성과 변화 경로를 분석하기 위한 기반을 마련한다(<그림 1>).

이러한 설정을 바탕으로, 본 연구는 시계열 구조를 활용하여 기간 간 매칭 구조를 구축하고, 서로 다른 시간 구간에서 소비자 행동과 결과 변수를 구분하여 측정한다. 구체적으로, 데이터를 연도 단위로 구분하고, 이전 기간(t 기)에서는 소비자의 장바구니 특성을 기반으로 행동 변수를 구성한다. 이후 다음 기간($t+1$ 기)에서는 소비자 가치와 관련된 결과 변수, 즉 구매 빈도, 소비 금액, 구매



<그림 1> 연구 모델

최근성 등을 측정한다. 실증 분석에서는 “2022년 행동-2023년 결과”, “2023년 행동-2024년 결과”와 같은 다수의 기간 간 매칭 구조를 구축한다. 또한 데이터 처리 과정에서는 인접한 두 기간 모두에서 소비 기록이 존재하는 소비자만을 유지하여 표본 간 매칭 가능성과 변수 간 대응 관계를 확보한다. 모든 설명 변수는 이전 기간에서 도출되며, 결과 변수는 이후 기간을 기준으로 계산된다. 이를 통해 시간적으로 명확한 선후 관계를 형성한다. 이러한 기간 간 분석 틀을 통해 본 연구는 “행동이 선행하고 결과가 후행한다”는 식별 논리를 시간 차원에서 구현하며, 이후 실증 분석을 위한 명확하고 견고한 기반을 제공한다.

2. 장바구니 스타일 군집 분석

2.1 장바구니 특성

2.1.1 구성(Composition) 특성

구성 특성은 장바구니의 수량 및 가격 측면에서의 기본 속성을 설명하며, 단일 구매 행위의 경제적 규모와 가격 구조를 반영한다. 본 연구는 다음의 세 가지 지표를 구성한다(<표 1>). 상품 수는 장바구니에 포함된 상품 기록의 개수를 의미하며, 단일 구매 행위의 규모를 측정한다.

<표 1> 장바구니 수준 특성의 정의 및 구성 방법

특성 구성	특성명	구성 방식
구성 특성	상품 수	장바구니 내 상품 수 집계
	총지출 로그값	상품별 (단가×수량) 합산 로그 변환 적용
	평균 단가	장바구니 내 상품 단가 평균 구매 수량 가중치 미적용
복잡성 특성	상품 분포 엔트로피	상품 분포 기반 엔트로피 계산 소비 분산 정도 측정
상황 특성	구매 시점	구매 시간대 (0~23시)
	주말 여부	주말(토/일)=1, 평일=0
	채널 유형	온라인=1, 오프라인=0

총지출 로그값은 먼저 장바구니 내 각 상품의 지출 금액을 합산하여 총지출을 계산한 후, 이에 로그 변환을 적용한다. 이는 극단값이 분석 결과에 미치는 영향을 완화하기 위함이다. 평균 단가는 장바구니 내 상품 기록에 해당하는 평균 상품 단가를 의미한다. 즉, 각 상품의 단가를 단순 평균하며 구매 수량에 따른 가중치는 적용하지 않는다. 이를 통해 장바구니의 가격 수준을 나타낸다. 이상의 지표는 “수량-지출-가격”의 세 가지 측면에서 장바구니의 기본적인 경제적 특성을 포착하며, 서로 다른 소비 강도와 가격 구조를 지닌 구매 행태를 구분하는 데 기초를 제공한다.

2.1.2 복잡성(Complexity) 특성

복잡성 특성은 장바구니 내부의 상품 조합 구조가 지니는 다양성과 분산 정도를 나타낸다. 이를 통해 “단일 집중형 장바구니”와 “다양 분산형 장바구니”를 구분하고, 쇼핑 구조의 복잡성을 파악할 수 있다. 상품 분포 엔트로피는 장바구니 내 상품 항목의 분포를 기반으로 엔트로피 지표를 계산한다. 장바구니 내 상품이 특정 항목에 집중되어 있는 경우(예: 소수 상품의 반복 구매)

에는 해당 값이 낮게 나타난다. 반면, 상품 종류가 보다 다양하고 분포가 균형을 이룰수록 해당 값은 높게 나타난다.

2.1.3 상황(Context) 특성

상황 특성은 쇼핑 행위가 발생한 시점의 시간적·채널적 상황을 설명하며, 장바구니가 놓인 외부 환경을 보완적으로 반영한다. 본 연구는 다음의 세 가지 변수를 구성한다. 구매 시점(시간대)은 쇼핑 행위가 발생한 시간(0~23시)을 의미하며, 하루 내 시간대별 소비 패턴의 차이를 반영한다. 주말 여부는 쇼핑이 주말(토요일 또는 일요일)에 발생한 경우 1, 그렇지 않은 경우 0으로 설정한다. 이를 통해 평일과 여가 시간대의 소비 행태를 구분한다. 채널 유형은 소매업체 코드를 기준으로 거래를 온라인과 오프라인 채널로 구분한다. 온라인은 1, 오프라인은 0으로 설정하며, 채널 환경에 따른 구매 행태의 차이를 반영한다.

2.2 군집 결과의 안정성 분석 및 K 값 선택

군집분석에서 적절한 군집 수(K)를 결정하는 것은 이후 결과 해석에 중요하다. 이에 본 연구는 기존의 군집 품질 지표에 더해 안정성 지표를 도입하여, 서로 다른 K 값에 따른 군집 결과를 종합적으로 평가한다. 구체적으로, 군집 품질은 실루엣 계수와 Calinski-Harabasz 지수를 통해 측정한다. 안정성은 다중 무작위 초기화 조건에서 산출된 조정 랜드 지수(ARI)와 정규화 상호정보량(NMI)의 평균과 표준편차를 기반으로 평가한다.

군집 품질만을 고려하는 접근과 달리, 안정성 지표는 서로 다른 초기화 조건에서도 군집 결과가 일관되게 나타나는지를 추가적으로 검토한다. 만약 초기화에 따라 결과의 변동이 크고(즉, 표준

편차가 높은 경우), 해당 군집 구조는 초기 조건에 민감하며 상대적으로 낮은 안정성을 가진다고 볼 수 있다. K=4부터 K=8까지의 비교 결과를 보면, 모든 모형에서 전반적인 일관성은 높은 수준으로 나타난다(ARI와 NMI의 평균값이 모두 1에 근접). 이는 데이터 자체가 비교적 명확한 구조를 가지고 있음을 시사한다. 그러나 보다 세밀하게 살펴보면, K 값에 따라 차이가 존재한다.

군집 품질 측면에서, 실루엣 계수는 K=4에서 가장 높은 값(.441)을 보이며, 이후 K가 증가할수록 점진적으로 감소한다. 이는 군집 간 구분도가 점차 약화됨을 의미한다. 반면, CH 지수는 K=5에서 최대값을 나타내지만 K=4와의 차이는 크지 않으며, 이후 K 값이 증가해도 지속적인 개선은 나타나지 않는다. 안정성 측면에서는 K=4와 K=5 모두 높은 수준을 보인다(ARI와 NMI 평균이 1에 근접하고 표준편차가 낮음). 이는 서로 다른 초기화 조건에서도 군집 결과가 안정적으로 재현됨을 의미한다. 그러나 K가 더 증가할수록 안정성은 감소하는 경향을 보인다. 예를 들어 K=8에서는 ARI와 NMI의 표준편차가 뚜렷하게 증가하여, 군집 결과가 초기 조건에 보다 민감해짐을 확인할 수 있다.

종합적으로 볼 때, K=5는 일부 지표에서 상대적으로 우수한 값을 보이지만, 실루엣 계수가 K=4보다 낮아 군집 구조의 명확성이 감소한다. 또한 더 큰 K 값에서는 안정성 저하 문제가 나타난다. 이에 비해 K=4는 군집 품질과 안정성 간의 균형이 가장 잘 유지되며, 군집 수가 비교적 적어 이후 변수 구성 및 경제적 해석에도 유리하다. 따라서 본 연구는 최적의 군집 수로 K=4를 선택한다(<표 2>).

2.3 군집분석 결과

2022~2023년의 거래 데이터를 기반으로, 본 연

<표 2> 서로 다른 군집 수(K)에 따른 군집 품질 및 안정성 지표

군집의 개수 (K)	응집도와 분리도 (Silhouette)	밀집도 (Calinski_Harabasz)	일치도 (ARI_mean)	안정성 (ARI_std)	유사성 (NMI_mean)
4	.441	59833.876	.993	.006	.988
5	.437	64860.194	.994	.007	.991
6	.403	64329.704	.986	.010	.981
7	.395	64678.950	.991	.009	.988
8	.369	63624.566	.978	.017	.977

구는 장바구니 수준에서 상품 수, 가격 수준, 그리고 상품 분포의 다양성 등의 특성을 추출하고, 군집분석을 통해 서로 뚜렷한 차이를 보이는 네 가지 장바구니 스타일을 도출하였다. 전반적으로 이러한 유형은 구매 규모, 가격 수준, 상품 분포의 다양성이라는 세 가지 차원을 중심으로 구분된다. 또한 이러한 차이를 바탕으로 각 유형은 구매 시간과 채널 측면에서도 비교적 안정적인 대응 관계를 보인다. 구체적으로 다음과 같다.

첫째, 저가 소량형 장바구니이다. 이 유형은 상품 수가 적고 가격 수준이 낮으며, 상품 분포의 다양성 또한 낮은 것이 특징이다. 이는 구매가 소수의 상품에 집중되어 있음을 의미한다. 이러한 특성 조합은 일상적인 보충 구매와 유사하다. 상황적으로는 주로 오전 시간대에 발생하며, 오프라인 채널 비중이 높다. 이는 일상적인 이동 경로 내에서 부수적으로 이루어지는 소규모·단순 구매 행태를 반영한다.

둘째, 고가 온라인 단일상품형 장바구니이다. 이 유형은 상품 수는 적지만 가격 수준이 매우 높고, 상품 분포 다양성은 가장 낮게 나타난다. 이는 구매가 단일 상품에 강하게 집중되어 있음을 의미한다. 해당 소비는 주로 온라인 채널에서 발생하며, 특히 새벽 시간대에 집중되는 경향을 보인다. 전반적으로 “단일 상품-고가격”의 특성을 가지며, 소비자가 명확한 수요에 기반하여 빠르게

구매를 완료하고, 다양한 상품 비교나 조합 선택은 상대적으로 적게 수행함을 시사한다.

셋째, 고다양성 대량구매형 장바구니이다. 이 유형은 세 가지 차원에서 또 다른 극단을 형성한다. 상품 수가 가장 많고, 상품 분포 다양성 또한 가장 높아 다양한 상품을 비교적 균형 있게 구매하는 특성을 보인다. 이는 조합적 구매 특성이 뚜렷한 유형이다. 일반적으로 이러한 특성은 집중적 구매 상황과 관련된다. 실제로 해당 장바구니는 주말에 더 자주 나타나며, 주로 오프라인 채널에서 발생한다. 이는 특정 시점에 소비자가 체계적으로 상품을 보충하는 행동 패턴을 반영한다.

넷째, 중간 규모 조합형 장바구니이다. 이 유형은 모든 차원에서 중간 수준을 보인다. 상품 수는 적절한 수준이며, 가격 수준도 중간 정도이고, 상품 분포 다양성 또한 일정 수준의 조합 특성을 나타낸다. 다른 유형에 비해 전반적으로 균형적인 특성을 가진다. 상황적으로는 주로 낮 시간대에 발생하며, 오프라인 채널 중심으로 나타난다. 이는 일상적인 쇼핑 과정에서 다양한 필요를 동시에 충족하는 일반적인 조합 구매 행태를 의미한다(<표 3>).

종합적으로, 이 네 가지 장바구니 스타일은 구매 규모, 가격 수준, 상품 분포 다양성 측면에서 명확하게 구분된다. 그러나 이들은 서로 완전히 분리된 소비 형태라기보다, 소비자가 다양한 상황

<표 3> 군집분석 결과

군집	상품 수	소비 금액	평균 가격	상품 카테고리 엔트로피	구매 시간대	주말 여부	채널 온라인 여부	군집 크기
0	1.501	8.946	6,643.377	.314	6.759	.288	.127	13,539
1	1.322	10.438	35,577.226	.181	2.085	.265	.937	13,457
2	11.637	11.307	7,032.825	2.399	7.114	.476	.019	7,678
3	4.615	10.356	7,706.666	1.431	6.934	.359	.024	14,017

에 따라 전환할 수 있는 여러 행동 패턴을 구성한다. 실제 데이터에서도 동일한 소비자는 특정 유형에 고정되기보다, 여러 유형의 장바구니 사이를 이동하는 경향을 보인다.

따라서 본 연구는 군집 결과를 정적인 소비자 분류로 해석하기보다, 개인 수준에서의 동적 조합 구조에 주목한다. 구체적으로, 특정 유형이 얼마나 지배적인지, 유형 간 전환 빈도가 어떠한지, 그리고 전체 분포의 분산 정도가 어떻게 나타나는지를 분석한다.

3. 모델 및 변수 설계

3.1 모델

소비자의 쇼핑 행동 구조가 이후 소비 행동에 미치는 영향을 분석하기 위해, 본 연구는 시차를 고려한 회귀모형을 구축하였다. 구체적으로, 이전 단계의 장바구니 행동 특성과 향후 소비 결과를 연결하여 분석한다. 종속변수는 향후 구매 빈도, 향후 소비 금액(로그값), 그리고 향후 구매 최근성으로 구성되며, 각각 소비 활동성, 소비 강도, 소비 지속성의 세 가지 측면에서 소비자 가치를 설명한다.

핵심 설명 변수는 쇼핑 행동의 동적 특성과 구조적 특성을 포함한다. 먼저, 장바구니 스타일 전환 횟수를 통해 소비자가 연속적인 구매 과정에

서 서로 다른 쇼핑 스타일 간에 전환하는 빈도를 측정한다. 또한, 장바구니 스타일 분포 엔트로피를 사용하여 소비자가 서로 다른 스타일에 어떻게 분포되어 있는지를 나타낸다. 아울러, 장바구니 주도적 스타일을 범주형 변수로 포함하여 서로 다른 소비 구조 간의 차이를 통제한다. 소비업체 다양성을 조절변수로 도입하여, 소비자가 이전 단계에서 접한 소매업체의 범위를 측정하고 소비업체의 다양성을 반영한다. 모형에서는 연도 고정 효과를 통제하며, 추정 과정에서는 강건 표준오차를 적용한다.

추가로, 변수 간 상관관계 및 다중공선성 문제를 점검하기 위해 상관분석과 분산팽창지수(VIF) 검정을 수행하였다. 그 결과, 대부분의 변수에서 VIF 값이 3 이하로 나타나 다중공선성 문제는 없다(<부록 표 1>). 일부 변수에서 상대적으로 높은 값이 관찰되었으나 일반적인 기준을 크게 초과하지 않아 분석 결과에 중대한 영향을 미치지 않는 것으로 판단된다. 관련 결과는 부록에 제시하였다.

3.2 변수 설계

3.2.1 독립변수

앞서 설정한 시간 구분을 바탕으로, 본 연구는 소비자 수준에서 쇼핑 스타일의 동적 특성을 추가로 분석한다. 구체적으로, 이전 기간 표본에서

소비자가 시간 순서에 따라 형성한 장바구니 시계열을 기반으로, 연속적인 구매 과정에서 각 장바구니 스타일이 어떻게 나타나는지를 추적한다. 이를 통해 서로 다른 유형 간 전환 양상과 전체 분포의 구조적 특성을 파악한다. 이러한 분석을 토대로, 본 연구는 쇼핑 스타일의 동적 변화를 반영하는 지표를 구성한다. 여기에는 특정 유형의 지배 정도, 유형 간 전환 빈도, 그리고 전체 분포의 다양성이 포함된다. 이러한 지표는 단일 구매나 전체 소비의 단순 집계가 아니라, 소비자가 여러 차례의 연속 구매에서 보이는 행동 궤적을 기반으로 산출된다. 따라서 소비 경로의 변화 과정을 보다 정확하게 반영할 수 있다.

또한, 해당 지표들은 모두 이전 기간의 행동 데이터를 기반으로 구성되며, 시간적으로 이후의 결과 변수에 선행한다. 이를 통해 잠재적인 역인과성 문제를 일정 부분 완화한다. 이와 같은 장바구니 시계열을 기반으로, 본 연구는 소비자 수준에서 쇼핑 스타일을 정량적으로 측정한다. 구체적으로, 소비자가 반복적인 구매 과정에서 “주로 어떤 장바구니 스타일을 보이는지”, “서로 다른 장바구니 스타일 간 전환이 발생하는지”, 그리고 “전체 분포가 집중되어 있는지 또는 분산되어 있는지”의 세 가지 측면을 중심으로 설명 변수를 구성한다. 각 변수의 정의는 <표 4>와 같다. 또한, 상호작용

항을 포함한 회귀모형에서 다중공선성 문제를 완화하고 해석의 안정성을 높이기 위해, 연속형 변수(전환 횟수, 분포 엔트로피)는 표준화하였다.

3.2.2 종속변수

앞서 구성한 독립변수를 바탕으로, 본 연구는 소비 결과의 관점에서 소비자의 후속 단계 행동을 추가로 분석한다. 구체적으로, “빈도, 금액, 최근성”의 기본 개념을 적용하여 구매 횟수, 소비 금액, 그리고 최근 구매 시점의 세 가지 측면에서 소비자의 소비 상태를 측정한다.

본 절에서는 다음 시점에서 실제로 나타난 소비 결과에 초점을 둔다. 연도 단위로 소비자의 구매 기록을 집계하여 해당 연도의 구매 빈도, 총 소비 금액, 그리고 연말 기준 최근 구매 시점을 산출한다. 이후 이러한 결과 변수를 시간 순서에 따라 이전 연도의 장바구니 스타일 특성과 연결함으로써 특정 연도에 형성된 소비자의 쇼핑 방식이 다음 연도의 소비 행동을 분석하는 데 활용된다(<표 5>).

3.2.3 조절변수

앞서 본 연구는 소비자가 쇼핑 과정에서 형성하

<표 4> 독립변수

변수명	구성 방식
장바구니 주도적 스타일	- 4개 스타일 기준 분류 (저가 소량형, 고가 온라인 단일형, 고다양성 대량형, 중간 규모 조합형) - 출현 빈도 기준 최빈 스타일 선택 - 더미 변수화(기준: 저가 소량형)
장바구니 스타일 전환 횟수	- 시계열 기준 정렬 - 인접 시점 스타일 변화 시 +1 - 전체 기간 누적
장바구니 스타일 분포 엔트로피	- 스타일별 비중 계산 - 엔트로피 산출

<표 5> 종속변수

변수명	구성 방식
구매 빈도	연도별 구매 횟수 집계
소비 금액	연도별 소비금액 합산
구매 최근성	연도 말 기준 마지막 구매 시점과의 시간 차이

는 행동 경로를 분석하였다. 예를 들어, 특정 쇼핑 방식을 지속적으로 유지하는지, 또는 서로 다른 쇼핑 방식 간에 빈번하게 전환하는지를 살펴보았다. 이러한 장바구니 스타일이 소비에 미치는 영향이 소비자에게 다르게 나타나는지 고민해 볼 필요가 있다. 각각의 소비자가 접하는 소비 환경은 서로 다르다. 일부 소비자는 제한된 수의 특정 소매업체에서 주로 소비를 수행하여 선택 범위가 상대적으로 좁다. 반면, 다른 소비자는 다양한 소매업체를 이용하며 보다 다양한 상품과 소비 상황에 노출된다. 이와 같은 차이로 인해, 유사한 쇼핑 행동을 보이더라도 그 의미는 달라질 수 있다. 예를 들어, 동일하게 서로 다른 장바구니 스타일 간 전환이 발생하더라도, 소비 범위가 제한된 소비자의 경우 이러한 변화는 유사한 상품과 고정된 구매 환경 내에서 이루어지며, 기존 선택지 간 반복적인 이동에 가까운 특성을 보인다. 반면, 다양한 소매업체를 이용하는 소비자의 경우 이러한 전환은 새로운 상품, 새로운 가격대, 또는 새로운

소비 상황을 동반할 가능성이 높으며, 보다 탐색적인 행동으로 해석될 수 있다. 따라서 장바구니 스타일 특성이 이후 소비에 미치는 영향은 소비자가 속한 소비 환경에 따라 달라질 가능성이 있다. 이러한 차이를 반영하기 위해, 본 연구는 “소매업체 다양성” 변수를 도입한다. 해당 변수는 일정 기간 동안 소비자가 이용한 서로 다른 소매업체의 수를 기준으로 구성되며, 소비 환경의 다양성을 측정한다.

IV. 연구 결과

1. 기술통계량

<표 6>은 주요 변수의 기술통계 결과를 제시한다. 전체적으로 볼 때, 소비자의 다음 단계 평균 구매 빈도는 17.736회이며, 평균 소비 금액(로그값)은 12.678, 평균 구매 간격은 약 204일로 나타난다. 이는 표본 내 소비자 간 소비 행동에 상당한 이질성이 존재함을 의미한다.

장바구니 스타일 전환 횟수의 평균은 12.561로 나타났으나, 표준편차가 18.051로 크게 나타났고 최대값은 151에 달한다. 이는 소비자 간 소비 경로의 변동 정도에 상당한 차이가 존재함을 보여

<표 6> 기술통계량

변수명	평균값	표준편차	최솟값	최댓값
장바구니 스타일 전환 횟수	12.561	18.051	0	151
장바구니 스타일 분포 엔트로피	.764	.408	.000	1.386
소비업체 다양성	1.597	.962	1	8
구매 빈도	17.736	30.488	1	335
소비 금액(로그값)	12.678	1.602	6.909	15.913
구매 최근성	204.453	129.140	-1	364
표본 수	643			

준다. 장바구니 스타일 분포 엔트로피는 전반적으로 변동 폭이 비교적 제한적이며, 이는 대부분의 소비자가 일정 수준의 다양성을 보이지만 서로 다른 스타일 간에 완전히 균등하게 분포되지는 않음을 시사한다.

또한, 소비업체 다양성의 평균은 1.597로 나타나, 대부분의 소비자는 소수의 특정 소매업체에 집중하여 소비를 수행하는 반면, 일부 소비자는 여러 소매업체에 걸쳐 소비하는 경향을 보인다.

장바구니 스타일 전환 횟수와 구매 빈도는 모두 일정 수준의 우측 편향 분포를 보인다. 이는 일부 고활동 소비자가 전체 결과에 상대적으로 큰 영향을 미칠 가능성을 의미한다. 이러한 잠재적 이분산 문제를 완화하기 위해, 이후 회귀분석에서는 모두 강건 표준오차를 적용하여 추정한다.

2. 결과

“저가 소량형 장바구니”를 기준 집단으로 설정한 결과, 서로 다른 주도 장바구니 스타일 간 소비 행동에 유의한 차이가 나타났다(<표 7>). 구체적으로, “고가 온라인 단일상품형 장바구니”를 주도적으로 보이는 소비자는 구매 최근성이 유의하게 높게 나타났다. 이는 재구매까지의 간격이 더 길다는 것을 의미한다. 이러한 결과는 단일 고가 상품 중심의 소비가 명확한 수요에 기반한 일회성 구매 성격을 가지며, 단기간 내 반복 구매로 이어지기 어렵다는 점을 시사한다.

또한, “고다양성 대량구매형 장바구니”를 주도적으로 보이는 소비자는 소비 금액에서 유의한 정(+의 영향을 보였다. 이는 집중적이며 다품목으로 구성된 구매가 전체 소비 금액을 증가시키

<표 7> 회귀분석 결과

변수	구매 최근성	구매 빈도	소비 금액
장바구니 주도적 스타일(고가 온라인 단일상품형)	38.551***	-1.400	.023
장바구니 주도적 스타일(고다양성 대량구매형)	-16.004	2.060	.779***
장바구니 주도적 스타일(중간 규모 조합형)	-16.968	7.278*	.716***
장바구니 스타일 전환 횟수	-0.998***	.773***	.025***
장바구니 스타일 분포 엔트로피	-4.607	-3.355	.084
소매업체 다양성	-11.497	-0.474	-0.063
고가 온라인 단일상품형×소매업체 다양성	-5.167	-0.170	.173
고다양성 대량구매형×소매업체 다양성	7.312	1.375	.172
중간 규모 조합형×소매업체 다양성	17.053	-2.498	.008
전환 횟수×소매업체 다양성	.115	-0.099	-0.005
분포 엔트로피×소매업체 다양성	5.854	.010	.157
상수항	190.985***	12.474***	12.234***
연도 고정효과	Yes	Yes	Yes
표본 수	643	643	643
Adj R ²	.123	.207	.118

* p<.10, *** p<.01 수준에서 통계적으로 유의하다.

는 데 기여한다.

한편, “중간 규모 조합형 장바구니”를 주도적으로 보이는 소비자는 소비 금액에서 유의한 정(+)의 영향을 보였으며, 구매 빈도에서도 한계적으로 유의한 정(+)의 영향을 나타냈다. 이는 단순 보충형 소비를 중심으로 하는 기준 집단과 비교할 때, 다양한 수요를 반영한 조합형 소비가 보다 안정적인 소비 패턴을 형성하고 전체 소비 수준을 높이는 데 기여함을 의미한다.

정태적인 주도 장바구니 스타일 외에도, 행동 경로의 동적 특성을 추가로 고려할 경우, 소비자가 서로 다른 장바구니 스타일 간에 전환하는 행위 역시 소비 행동에 중요한 영향을 미친다. 장바구니 스타일 전환 횟수는 세 가지 모형 모두에서 유의한 영향을 보였다. 구체적으로, 구매 빈도와 소비 금액에는 유의한 정(+)의 영향을 미치는 반면, 구매 최근성에는 유의한 부(-)의 영향을 나타냈다. 이는 장바구니 스타일 간 전환이 빈번할수록 소비자의 소비 활동성이 높고, 소비 규모가 크며, 재구매 간격이 짧아지는 경향이 있음을 의미한다.

앞서의 군집 결과를 고려할 때, 이러한 “전환 행동”은 소비자가 서로 다른 소비 상황에 따라 행동을 조정하는 과정으로 해석할 수 있다. 예를 들어, 일상적 보충 구매, 집중 구매, 단일 고가 상품 구매 등 다양한 소비 방식 간 전환을 통해, 소비자는 다양한 수요 상황에서 지속적으로 소비를 발생시킬 수 있으며, 이는 전체 소비의 향상으로 이어진다. 따라서 단일한 주도 장바구니 스타일만을 고려하는 것보다, 행동 경로의 동적 변화는 소비자 행동을 이해하는 데 정보를 제공한다.

반면, 장바구니 스타일 분포 엔트로피는 모든 모형에서 유의한 영향을 보이지 않았다. 이는 단순히 전체 분포의 균형성만으로는 소비의 차이를 충분히 설명하기 어렵다는 점을 시사한다.

종합하면, 주도 장바구니 스타일이 반영하는 정태적 소비 구조는 소비자의 기본적인 소비 성향을 설명하며, 이에 더해 서로 다른 장바구니 스타일 간의 동적 전환 행동은 소비 행동의 차이를 추가적으로 설명한다. 두 요소는 각각 “어떤 소비 방식을 선호하는가”와 “서로 다른 소비 방식 간 어떻게 조정하는 가”라는 두 가지 측면에서 소비자 행동을 설명한다.

결과 변수 측면에서 이러한 메커니즘은 구매 최근성, 구매 빈도, 소비 금액 전반에 걸쳐 나타나며, 각각 재구매 속도, 소비 활동 수준, 전체 소비 규모와 대응된다.

3. 조절효과

소매업체 다양성은 모든 모형에서 유의한 직접 효과를 보이지 않았다. 이는 소비자가 이용하는 소매업체의 수가 증가하더라도 구매 빈도나 소비 금액이 직접적으로 증가하지 않음을 의미한다.

장바구니 주도적 스타일, 스타일 전환 횟수, 분포 엔트로피와 소매업체 다양성 간의 상호작용항은 세 종속변수 모형 모두에서 유의하지 않은 것으로 나타났다. 즉, 소매업체 다양성은 쇼핑 장바구니 구조가 소비자 행동에 미치는 영향을 유익하게 변화시키지 않았다. 이 결과는 쇼핑 장바구니 구조가 소비자 행동에 미치는 영향이 소비 환경이 달라져도 비교적 안정적인 모습을 보여준다. 다시 말해, 소비자가 소수의 소매업체에 집중하든 여러 소매업체에 분산하든, 그 행동 패턴은 구매 빈도와 소비 금액에 유사하게 작용한다.

쇼핑 장바구니 스타일 전환 횟수는 소비 행동에 유의한 영향을 미친다. 그러나 이 효과는 소매업체 다양성에 따라 달라지지 않았다. 이는 소비자 행동이 외부 채널의 분산 정도보다 내부 행동 경로에 더 크게 영향을 받음을 의미한다. 또한 쇼핑

장바구니 스타일 분포 엔트로피와 관련된 상호작용항도 유의하지 않았다. 이는 소비 구조의 균형성이 소비 환경의 확장에 따라 추가적인 영향을 받지 않음을 보여준다.

V. 결론

1. 연구의 요약

본 연구는 장바구니 수준의 행동 데이터를 기반으로, “장바구니 주도적 스타일”과 “장바구니 스타일 간 전환 행동”이라는 두 가지 차원에서 소비자의 소비 행동을 체계적으로 분석하였다. 분석 결과, 서로 다른 주도 장바구니 스타일 간에는 소비 행동에서 뚜렷한 차이가 나타났으며, 동시에 소비자가 서로 다른 장바구니 스타일 간에 전환하는 행동 또한 구매 빈도, 소비 금액, 재구매 간격에 중요한 영향을 미치는 것으로 확인되었다.

구체적으로, 장바구니 주도적 스타일은 소비자가 “어떠한 방식으로 소비하는 경향이 있는지”를 설명한다. 예를 들어, 고가 온라인 단일상품형 장바구니를 주로 보이는 소비자는 명확한 수요에 기반한 일회성 구매를 선호하는 경향이 있으며, 이에 따라 재구매 간격이 더 길게 나타난다. 반면, 고다양성 대량구매형 또는 중간 규모 조합형 장바구니를 주로 보이는 소비자는 다품목 조합 구매를 수행하는 경향이 강하며, 이는 더 높은 총 소비 금액과 보다 안정적인 소비 패턴으로 이어진다. 이러한 결과는 소비자의 소비 구조 자체가 소비 방식과 소비 강도의 차이를 일정 부분 설명할 수 있음을 의미한다.

이와 함께, 동적 관점을 도입하면 장바구니 스타일 전환 횟수가 모든 모형에서 유의한 영향을 미치는 것을 확인할 수 있다. 전환이 빈번할수록

소비자의 구매 빈도와 소비 금액은 증가하고, 재구매 간격은 짧아지는 경향을 보인다. 이는 소비자가 특정 소비 방식에 고정되기보다는, 서로 다른 소비 상황에 따라 행동을 조정하고 있음을 시사한다. 예를 들어, 일상적인 보충 구매, 집중 구매, 단일 상품 구매 간 전환을 통해 다양한 수요 상황에서 지속적으로 소비를 발생시키며, 그 결과 전체적인 소비 행동이 향상된다.

한편, 장바구니 스타일 분포 엔트로피는 유의한 영향을 보이지 않았다. 이는 단순히 “소비가 얼마나 균등하게 분포되어 있는가”라는 관점만으로는 소비 차이를 충분히 설명하기 어렵다는 점을 의미한다. 이에 비해, 서로 다른 유형 간 전환이 실제로 발생하는지 여부와 같은 동적 행동은 소비 과정을 보다 잘 반영하는 것으로 나타난다.

또한, 조절효과 분석 결과에 따르면 소비 환경 다양성은 장바구니 구조와 소비 행동 간의 관계를 유의하게 조절하지 않는 것으로 나타났다. 즉, 소비자가 이용하는 소매업체의 범위와 관계없이, 장바구니 구조가 소비 행동에 미치는 영향은 비교적 안정적으로 유지되는 경향을 보인다.

종합하면, 소비자 행동은 장기적으로 형성된 소비 성향(장바구니 주도적 스타일)과 구매 과정에서 나타나는 행동 조정 방식(장바구니 전환 행동)이라는 두 가지 차원에서 이해될 수 있다. 다만, 이러한 행동 경로는 소비 환경에 의해 추가적으로 강화되거나 약화되기 보다는, 비교적 독립적으로 작용하는 경향을 보인다.

2. 학문적 시사점

본 연구의 시사점은 다음과 같다. 첫째, 본 연구는 장바구니 수준에서 소비자 행동을 분석함으로써 연구 관점을 “단일 구매 결과”에서 “구매 방식”으로 확장하였다. 분석 결과, 장바구니 스타일

에 따라 소비 행동에 연결되는 것으로 나타났으며, 이는 소비자 사이의 다른 특성 차이를 이해하는 데 도움을 준다.

둘째, 본 연구는 동적 관점을 도입하여 “장바구니 주도적 스타일”과 “장바구니 스타일 간 전환 행동”을 구분하였다. 분석 결과, 정태적인 소비 성향보다 서로 다른 쇼핑 방식 간 전환이 소비 행동의 차이를 더욱 효과적으로 설명하는 것으로 나타났다. 이는 소비자 행동에서 “변화 과정”이 갖는 중요성을 강조하는 결과이다.

셋째, 본 연구는 소비자의 행동을 소비 업체 다양성과 결합하여 분석하였다. 그 결과, 소비 업체가 다양한지는 소비 결과에 직접적인 영향을 미치지 않았다. 그리고 소비 행동이 결과로 이어지는 과정에서도 조절 효과가 유의하지 않았다. 이러한 결과는 소비자 행동과 소비 업체 다양성의 관계가 제한적이며, 소비 행동이 외부 환경보다는 내부 행동 패턴에 의해 형성된다는 점을 보여준다.

3. 실무적 시사점

본 연구의 실무적 시사점은 다음과 같다. 첫째, 결과는 소비자 가치의 형성이 단순히 소비 수준에만 의존하는 것이 아니라, 서로 다른 쇼핑 방식 간 전환과도 밀접하게 관련되어 있음을 보여준다. 단순히 구매 횟수나 소비 금액을 증가시키는 것보다, 소비자가 다양한 쇼핑 방식 간에 전환하도록 유도하는 것이 보다 효과적인 가치 제고 전략이 될 수 있다. 실제 운영 관점에서, 우선 소비자의 장바구니 주도적 스타일을 기준으로 보다 정교한 세분화가 가능하다. 분석 결과, 각 유형은 서로 다른 소비 패턴과 대응된다. 예를 들어, 고가 온라인 단일상품형 장바구니를 주로 보이는 소비자는 수요가 명확하고 의사결정이 빠르지만, 소비

가 비교적 분산되어 있으며 재구매 간격이 긴 경향이 있다. 반면, 중간 규모 조합형 또는 고다양성 대량구매형 장바구니를 주로 보이는 소비자는 보다 지속적이고 안정적인 소비를 형성하며, 높은 총 소비 금액으로 이어지는 경향이 있다. 이에 따라 운영 측면에서는 유형별로 차별화된 전략을 적용할 수 있다. 예를 들어, 단일상품형 소비자에게는 재구매 주기를 단축하기 위해 보충 알림이나 유사 상품 추천을 제공하고, 조합형 소비자에게는 상품 묶음 추천이나 멤버십 프로그램을 통해 장기적 가치를 강화할 수 있다.

셋째, 소매업체가 “소비 결과의 증대”에서 “소비 방식의 유도”로 전략의 초점을 전환해야 한다는 점이다. 서로 다른 장바구니 스타일 간 전환이 빈번한 소비자는 더 높은 구매 빈도와 소비 금액을 보이는 경향이 있다. 이는 쇼핑 방식 간 전환 자체가 소비자 가치 향상의 핵심 메커니즘임을 의미한다. 실제로는 연속적인 상품 추천과 진열 방식을 통해 이러한 전환을 유도할 수 있다. 예를 들어, 단일상품 구매 이후 다상품 조합 구매로 확장하도록 유도하거나, 일상적인 보충 구매 상황에서 관련 상품 또는 세트 상품을 제안하여 구매 범위를 점진적으로 확대할 수 있다.

마지막으로, 본 연구는 소비자 관리가 단순한 소비 금액이나 빈도와 같은 결과 지표에 머무르지 않고, 행동 과정 자체에 대한 이해로 확장될 필요가 있음을 시사한다. 단순히 소비자의 활동성을 높이는 것보다, 현재 어떤 쇼핑 방식을 보이고 있는지, 그리고 다른 방식으로 전환할 가능성이 있는지를 파악하는 것이 중요하다. 예를 들어, 특정 쇼핑 방식에 장기간 머무르는 소비자에게는 추천 시스템을 통해 행동 패턴의 전환을 유도할 수 있으며, 이미 다양한 쇼핑 방식 간 전환을 보이는 소비자에게는 이러한 행동을 강화하여 보다 안정적인 소비 순환을 형성하도록 지원할 수 있

다. 이처럼 운영의 초점을 “단일 소비의 촉진”에서 “소비 방식의 전환 유도”로 전환할 때, 보다 지속적인 소비자 가치 증대가 가능해진다.

4. 한계점 및 향후 연구 방향

본 연구는 거래 행동 데이터에 기반하여 분석을 수행하였기 때문에, 소비자의 개인적 특성(persona)을 직접적으로 반영하지 못하였으며, 동일한 장바구니 구조가 서로 다른 소비자에게서 어떠한 맥락에서 나타나는지를 구분하는 데에는 한계가 있다. 또한 연 단위로 소비자 행동을 구성함에 따라, 프로모션이나 계절적 요인 등 단기적 요인이 장바구니 구조 및 전환 과정에 미치는 영향을 충분히 반영하지 못하였다.

방법론적 설계 측면에서, 본 연구는 (t-1) 시점의 시간 지연 구조를 도입함으로써 변수 간 시간적 선후관계를 구분하고, 이에 따라 잠재적인 동시성(simultaneity) 및 내생성 문제를 완화하고자 하였다. 그러나 이러한 접근은 여전히 관측 데이터에 기반하고 있기에, 내생성으로 인한 영향을 근본적으로 완전히 제거하기에는 한계가 있다. 따라서 본 연구의 결과는 시간적 순서를 통제한 상태에서 연관성에 대한 증거로 해석하는 것이 타당하며, 엄밀한 인과관계로 이해하기에는 신중한 접근이 필요하다.

나아가, 본 연구는 관찰된 행동 패턴을 중심으로 분석을 수행하였으나, 이러한 전환 행동이 어떠한 메커니즘에 의해 형성되는지까지는 구분하지 못한다는 한계가 존재한다.

향후 연구에서는 소비자 개인 특성과 거래 데이터를 결합하여 행동 이질성을 보다 정교하게 분석할 필요가 있다. 보다 세분화된 시간 단위 또는 상품 수준의 데이터를 활용하여 소비자 행동 경로의 변화를 동적으로 탐색할 수 있을 것이다. 또

한 다양한 상황에서 쇼핑 방식 전환이 어떻게 형성되는지를 규명하면, 소비자 행동에 대한 보다 심층적인 이해를 제공할 수 있을 것이다.

논문접수일: 2026. 01. 29.

1차 수정본 접수일: 2026. 03. 30.

게재확정일: 2026. 04. 14.

이해 상충에 관한 보고

본 논문과 관련된 잠재적 이해 상충 관계가 없음을 보고함.

연구비 지원

본 논문은 어떠한 연구비 지원도 받지 않음.

감사의 글

이 논문은 2025년 한국유통학회와 텐큐브의 학술데이터지원사업 지원을 받아 수행된 연구임.

연구 데이터 접근 가능성

본 연구에 사용된 데이터는 기밀 유지로 인해 제공될 수 없음.

저자 기여 항목

연구개념화: 황자성.

데이터 큐레이션/조사: 김태완.

데이터 분석/검증: 황자성.

방법론: 황자성, 김태완.

원고 초안 작성: 황자성, 이로.

원고 검토 및 편집: 황자성, 이로, 김태완.
자금 조달/자원 확보: 김태완.

윤리 심의 승인에 관한 보고

본 연구는 인간 및 동물 참여자가 없으므로 IRB/IACUC 심의가 필요하지 않음.

생성형 AI 사용에 관한 선언

본 논문은 생성형 AI의 사용과 무관함.

참고문헌

- Abbasimehr, H. & Bahrini, A. (2022). An analytical framework based on the recency, frequency, and monetary model and time series clustering techniques for dynamic segmentation. *Expert Systems with Applications*, 192, 116373.
- Agrawal, R., Imielinski, T., & Swami, A. (1993). Mining association rules between sets of items in large databases. *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data* (pp. 207-216).
- Bai, T., Nie, J. Y., Zhao, W. X., Zhu, Y., Du, P., & Wen, J. R. (2018). An attribute-aware neural attentive model for next basket recommendation. *Proceeding of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval* (pp. 1201-1204).
- Bell, D. R., Corsten, D., & Knox, G. (2011). From point of purchase to path to purchase: How preshopping factors drive unplanned buying. *Journal of Marketing*, 75(1), 31-45.
- Bolton, R. N., Lemon, K. N., & Verhoef, P. C. (2004). The theoretical underpinnings of customer asset management: A framework and propositions for future research. *Journal of the Academy of Marketing Science*, 32(3), 271-292.
- Boztug, Y. & Reutterer, T. (2008). A combined approach for segment-specific market basket analysis. *European Journal of Operational Research*, 187(1), 294-312.
- Brijs, T., Swinnen, G., Vanhoof, K., & Wets, G. (1999). Using association rules for product assortment decisions: A case study. *Proceedings of the fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 254-260).
- Engel, J. F., Blackwell, R. D., & Miniard, P. W. (1986). *Consumer behavior*. Dryden Press.
- Fader, P. S., Hardie, B. G. S., & Ka Lok Lee. (2005). RFM and CLV: Using iso-value curves for customer base analysis. *Journal of Marketing Research*, 42(4), 415-430.
- Griva, A., Bardaki, C., Pramataris, K., & Papakiriakopoulos, D. (2018). Retail business analytics: Customer visit segmentation using market basket data. *Expert Systems with Applications*, 100, 1-16.
- Guidotti, R., Coscia, M., Pedreschi, D., & Pennacchioli, D. (2015). Behavioral entropy and profitability in retail. *Proceeding of the 2015 IEEE International Conference on*

- Data Science and Advanced Analytics (DSAA)* (pp. 1-10).
- Gupta, S., Lehmann, D. R., & Stuart, J. A. (2004). Valuing customers. *Journal of Marketing Research*, 41(1), 7-18.
- Hidasi, B., Karatzoglou, A., Baltrunas, L., & Tikk, D. (2015). Session-based recommendations with recurrent neural networks. *arXiv preprint arXiv:1511.06939*.
- Hui, S. K., Inman, J. J., Huang, Y., & Suher, J. (2013). The effect of in-store travel distance on unplanned spending: Applications to mobile promotion strategies. *Journal of Marketing*, 77(2), 1-16.
- Jannach, D. & Adomavicius, G. (2016). Recommendations with a purpose. *Proceedings of the 10th ACM Conference on Recommender Systems* (pp. 7-10).
- Kotler, P. & Keller, K. L. (2016). *Marketing management*. Pearson.
- Lemon, K. N. & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. *Journal of Marketing*, 80(6), 69-96.
- Li, H. & Kannan, P. K. (2014). Attributing conversions in a multichannel online marketing environment: An empirical model and a field experiment. *Journal of Marketing Research*, 51(1), 40-56.
- McAlister, L. & Pessemier, E. (1982). Variety seeking behavior: An interdisciplinary review. *Journal of Consumer Research*, 9(3), 311-322.
- Neslin, S. A., Grewal, D., Leghorn, R., Shankar, V., Teerling, M. L., Thomas, J. S., & Verhoef, P. C. (2006). Challenges and opportunities in multichannel customer management. *Journal of Service Research*, 9(2), 95-112.
- Oliver, R. L. (1999). Whence consumer loyalty? *Journal of Marketing*, 63(4_suppl1), 33-44.
- Reicheld, F. F. & Sasser Jr, W. E. (2003). Quality comes to services. *Operations Management*, 105, 289.
- Reinartz, W. J. & Kumar, V. (2003). The impact of customer relationship characteristics on profitable lifetime duration. *Journal of Marketing*, 67(1), 77-99.
- Rendle, S., Freudenthaler, C., & Schmidt-Thieme, L. (2010). Factorizing personalized markov chains for next-basket recommendation. *Proceedings of the 19th International Conference on World Wide Web* (pp. 811-820).
- Rust, R. T., Zeithaml, V. A., & Lemon, K. N. (2004). Customer-centered brand management. *Harvard Business Review*, 82(9), 110-120.
- Tan, Y. K., Xu, X., & Liu, Y. (2016). Improved recurrent neural networks for session-based recommendations. *Proceedings of the 1st Workshop on Deep Learning for Recommender Systems* (pp. 17-22).
- Trusov, M., Bucklin, R. E., & Pauwels, K. (2009). Effects of word-of-mouth versus traditional marketing: Findings from an internet social networking site. *Journal of Marketing*, 73(5), 90-102.
- Tsai, C. Y. & Chiu, C. C. (2004). A purchase-based market segmentation methodology. *Expert Systems with Applications*, 27(2),

- 265-276.
- Verhoef, P. C. (2003). Understanding the effect of customer relationship management efforts on customer retention and customer share development. *Journal of Marketing*, 67(4), 30-45.
- Verhoef, P. C., Franses, P. H., & Hoekstra, J. C. (2002). The effect of relational constructs on customer referrals and number of services purchased from a multiservice provider: Does age of relationship matter? *Journal of the Academy of Marketing Science*, 30(3), 202-216.
- Wang, P., Guo, J., Lan, Y., Xu, J., Wan, S., & Cheng, X. (2015). Learning hierarchical representation model for nextbasket recommendation. *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 403-412).
- Wang, S., Sun, L., & Yu, Y. (2024). A dynamic customer segmentation approach by combining LRFMS and multivariate time series clustering. *Scientific Reports*, 14(1), 17491.
- Wedel, M. & Kamakura, W. A. (2000). *Market segmentation: Conceptual and methodological foundations*. Springer Science & Business Media.
- Zeng, W., Liu, J., & Zhang, B. (2024). Hierarchical representation learning for next basket recommendation. *Array*, 23, 100354.

<부록>

<부록 표 1> 분산팽창지수(VIF) 결과

변수	VIF
장바구니 주도적 스타일(고가 온라인 단일상품형)	1.614
장바구니 주도적 스타일(고다양성 대량구매형)	1.479
장바구니 주도적 스타일(중간 규모 조합형)	1.601
장바구니 스타일 전환 횟수	1.523
장바구니 스타일 분포 엔트로피	1.522
소매업체 다양성	5.370
고가 온라인 단일상품형×다양성	2.850
고다양성 대량구매형×다양성	1.446
중간 규모 조합형×다양성	2.626
전환 횟수×소매업체 다양성	2.174
분포 엔트로피×소매업체 다양성	1.901
연도 고정효과(2023)	1.222

The Impact of Shopping Cart Structure and Conversion Behavior on Consumer Value*

Zixing Huang**, Lu Li***, Taewan Kim****

ABSTRACT

Purpose: This study examines how basket structure and its evolution influence consumer consumption outcomes based on transaction data at the basket level. Unlike prior research that primarily focuses on single purchase outcomes, this study identifies four representative basket types through cluster analysis.

Research design, data, and methodology: At the individual level, it constructs dominant basket style and basket style switching behavior to capture both relatively stable consumption tendencies and dynamic behavioral changes. In addition, the study introduces consumption environment diversity—defined as the range of retailers and channels a consumer engages with—to examine its moderating role.

Results: The empirical results are as follows. First, consumers whose dominant basket style is diversified or combination-oriented exhibit significantly higher spending levels. Second, basket style switching frequency has a significant effect across all three outcome dimensions, leading to higher purchase frequency, greater spending, and shorter inter-purchase intervals.

Conclusions: Overall, the findings suggest that consumer performance is not only determined by consumption levels but is also closely related to the ability to switch between different basket structures. By focusing on “how baskets are constructed” and “how baskets evolve,” this study provides a new analytical framework for understanding consumer behavior. It also offers practical managerial implications for retailers seeking to optimize customer management strategies in an omni-channel environment.

Keywords: Basket-Based Consumer Behavior, Cluster Analysis (K-Means), Customer Value

* This study was conducted with the support of the 2025 Academic Data Support Program provided by the Korea Distribution Association (KODIA) and Tenqube.

** Ph. D. Candidate, Department of Business Administration, Konkuk University, First Author

*** Ph. D. Student, Department of Business Administration, Konkuk University, Co-Author

**** Professor, Department of Business Administration, Konkuk University, Corresponding Author