



ISSN: 2288-7709

JEMM website: <https://acom.s.kisti.re.kr/jemm>doi: <http://dx.doi.org/10.20482/jemm.2023.11.5.45>

유통업체 연구동향 분석: 백화점을 중심으로

Research Trend Analysis of the Retail Industry: Focusing on the Department Store

Hoe-Chang YANG¹

Received: September 22, 2023. Revised: September 26, 2023. Accepted: September 30, 2023.

Abstract

Purpose: As one of the continuous studies on the offline distribution industry, the purpose of this study is to find ways for offline stores to respond to the growth of online shopping by identifying research trends on department stores. **Research design, data and methodology:** To this end, this study conducted word frequency analysis, word co-occurrence frequency analysis, BERTopic, LDA, and dynamic topic modeling using Python 3.7 on a total of 551 English abstracts searched with the keyword 'department store' in scienceON as of October 10, 2022. **Results:** The results of word frequency analysis and co-occurrence frequency analysis revealed that research related to department stores frequently focuses on factors such as customers, consumers, products, satisfaction, services, and quality. BERTopic and LDA analyses identified five topics, including 'store image,' with 'shopping information' showing relatively high interest, while 'sales systems' were observed to have relatively lower interest. **Conclusions:** Based on the results of this study, it was concluded that research related to department stores has so far been conducted in a limited scope, and it is insufficient to provide clues for department stores to secure competitiveness against online platforms. Therefore, it is suggested that additional research be conducted on topics such as the true role of department stores in the retail industry, consumer reinterpretation, customer value and lifetime value, department stores as future retail spaces, ethical management, and transparent ESG management.

Keywords : Research Trend, Research Trend Analysis, Retail Industry, Department Store, Topic Modeling.

JEL Classification Code: C40, C46, C60, L10, M10.

1. Introduction

소비자의 소비행태와 관련되어 영향을 미치는 요인은 다양하다. 이는 소비자의 쇼핑행위에는 유형의 제품을 획득하는 것뿐

아니라 서비스를 포함한 레이아웃, 청결 및 분위기와 같은 다양한 경험적 속성이 영향을 미치기 때문이다(Dennis et al, 2002). Youn and Kim(2017)에 따르면 유통업은 다음과 같이 분류할 수 있다. 첫째, '무엇을 판매하는가(What to Sell)와 관련되어 직업이나

*본 연구는 IFDC 2023 에서 요약본으로 발표한 것으로, 이 논문은 장안대학교 2023 년도 연구비 지원에 의하여 연구되었음.

¹ First Author. Ph.D., Assistant Professor,
Department of Distribution Management, Jangan University,
South Korea. Email: pricezzang@jangan.ac.kr

© Copyright: The Author(s)
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

영업의 종류로 분류되는 업종(types of business)이 그 하나이다. 둘째, '어떠한 방법으로 판매하는가(How to Sell)와 관련되어 소비자 구매행동과 관련된 판매시스템으로서의 소매점 유형인 업태(types of operation)로의 구분을 들 수 있다. 이런 구분은 소비자의 관심 방향과도 같은 맥락에서 분류될 수 있다.

업태라는 용어의 사용은 1990년대 초반부터인 것으로 알려져 있다. 일본의 공정거래위원회는 업태를 '영업 형태화로 지적하고 점포의 형태, 거래상품의 종류, 판매정책, 일반소비자의 구매의식의 차이'라고 정의하고 있으며, 일본의 소매업협회는 '업태란 소비자의 구매관습의 변화에 대응한 소매업자의 영업형태이고, 업종과는 다르게 소비자 지향적인 분류'라고 정의하고 있다(Youn & Kim, 2017). 이와 같은 정의를 유통산업발전법에 대입해보면 대규모점포, 준대규모점포 등을 업태의 한 유형으로 볼 수 있다. 즉, 업태란 고객 세분화에 따른 소매기업간의 경쟁유발의 결과 생성된 소매업자들의 상품의 결합체가 여러 가지 마케팅 포인트와 결합하여 새로운 유형을 만들어 낸 것이다.

본 연구가 유통업태에 주목하는 이유는 1996년 유통시장 개방 후 선진화된 다양한 업태의 출현을 들 수 있다. 또한, 1인 가구 증가, 디지털기술 발달, 코로나 19와 같은 다양한 요인으로 인한 비대면 경제로의 전환의 혜택으로 등장한 온라인 유통채널로 인해 유통업태의 경쟁관계가 상당히 복잡해졌기 때문이다. 특히, 소비자들의 라이프스타일과 소비방식의 변화에 대응하기 위해 오프라인 유통업태는 O2O(online to Offline) 사업 확대 등 시스템 확충 등 다양한 영역에서의 변화를 꾀하고 있다. 연구자들 또한 유통플랫폼(Griffis et al., 2012; Groß, 2015; Holmes et al., 2014; Shang et al., 2017), 유통제품 혁신(Gao & Ko, 2017; Jiang et al., 2021), 신 유통전략(Kollmann et al., 2012; Li & Kannan, 2014), 전략적 배송(Nguyen et al., 2019; Toker et al., 2020) 등 트렌드 변화와 관련된 연구결과를 보고하며 오프라인 유통업태의 성장과 발전을 위한 아이디어를 제공하고 있다(Yang, 2022).

본 연구에서는 오프라인 유통업태에 대한 시리즈 연구 중 하나로 백화점에 대한 연구동향을 확인하고자 한다. 우리나라 백화점은 도시 중심에 입지하고 있고, 생활밀착형으로 운영되며, 다양한 상품 구색과 낮은 직매입 비중이라는 특징을 갖고 있다. 이런 이유로 백화점은 근거리·소량구매를 선호하는 최근 소비트렌드 뿐 아니라 타 업태 특히, 최근 성장세가 상당한 온라인쇼핑 분야에 비해 보유하고 있는 매장 이미지, 레이아웃, 청결 및 분위기에 우위를 점하고 있는 대표적인 유통업태 중 하나라고 판단된다.

본 연구의 목적은 오프라인 유통업태 중 백화점에 대한 그간 연구동향을 파악함으로써 오프라인매장이 온라인쇼핑의 성장세에

대응할 수 있는 방안을 모색하는 것이다. 이를 위해 본 연구에서는 scienceON에 등록된 논문의 영문초록을 Python 3.7을 이용하여 BERTopic과 LDA를 통해 연구되는 토픽을 분류하고 어떤 토픽이 연구자들의 관심주제인지를 확인함으로써 향후 백화점 발전을 위한 방향을 제시하고자 하였다.

본 연구결과는 오프라인 유통업태인 백화점이 비대면 경제로의 전환이라는 큰 물결을 이겨낼 수 있는 방안과 함께 오프라인 유통업태들에게도 상생과 발전을 위한 통찰력을 제공해줄 것으로 기대하였다. 또한 토픽모델링에서 주로 이용되던 LDA보다 더 정교하다고 인정받고 있는 BERTopic을 적용함으로써 장점과 한계점을 확인할 수 있을 것으로 기대하였다.

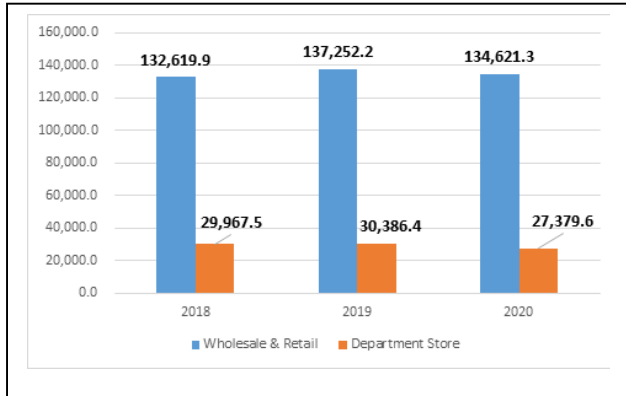
2. Theoretical Background

2.1. Department Store

표준산업분류에서 도·소매업으로 분류되는 유통산업은 2020년 기준 134.5조 원으로 전체 GDP의 7.3%를 차지하고 있으며, 2019년 기준으로 전체 사업체 수는 102.8만 개로 24.6%, 종사자 수는 323.9만 명으로 14.5%를 차지하고 있는 국가경제의 성장동력원 중 하나로 발돋움하였다(Korcham, 2021). 산업통상자원부의 보도자료에 따르면 백화점은 2020년 기준 매출액 27.4조 원으로 오프라인 전체 매출의 15.2%를 차지하고 있으며 2019년 대비 약 9.9%의 매출 감소가 있었지만 최근 코로나19 영향에 대한 기저효과와 잠재된 소비심리의 표출에 따라 매출이 폭발적으로 증가하고 있는 것으로 나타났다(MOTIE, 2022).

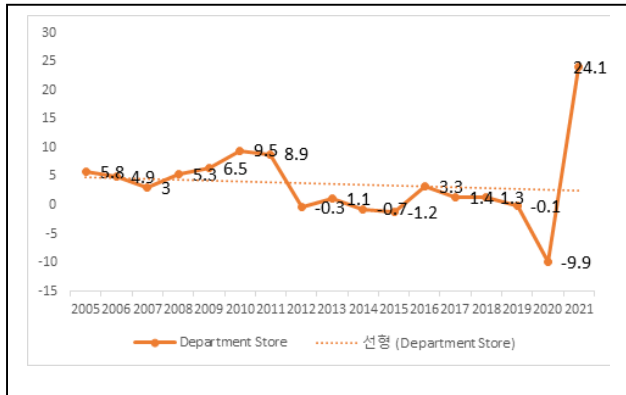
대한상공회의소의 보고자료에 따르면 산업별 국내총생산(GDP) 중 도·소매 전체 매출액은 2018년 132조 6,199억 원에서 2019년 137조 2,522억 원으로 약 4조 6,323억 원 증가하였다가 2020년에는 134조 6,213억 원으로 감소하였으며, 백화점은 2018년 29조 9,675억 원(22.6%)에서 2020년 27조 3,796억 원으로 지속적인 감소를 하는 것으로 나타났다(<Figure 1> 참조). 한편 백화점 매출 증감률(<Figure 2> 참조)에서 보는 바와 같이 2012년, 2014년, 2015년뿐 아니라 코로나19가 창궐한 2018년에도 -0.1%p 감소하였다가 2020년에는 -9.9%p로 대폭 감소하는 어려움을 겪었던 것이 사실이다. 이런 이유는 앞에서 언급한 바와 같이 감염병으로 인한 영업점 일시 폐쇄 등과 관련된 소비 위축, 소비자들의 비대면 소비 활성화와 같은 요인들이 영향을 미친 것으로 판단된다. 그러나 2021년에는 코로나19로 인해 그동안 억제되었던 소비심리가 표출되고 해외여행의 제약으로 인한 유명브랜드에 대한 매출 확대와 아동 스포츠부문을 중심으로 전 품목의 매출이 상승하면서 약 24.1% 매출증가를 달성한 것

으로 보고되고 있다(MOTIE, 2022).



Source) Diagrammed the data from MOTIE (2022)

Figure 1: Comparison of wholesale and retail gross domestic product (GDP) and department store sales from 2018 to 2020



Source) Diagrammed the data from MOTIE (2022)

Figure 2: Department store sales change rate by year

세계 최초의 백화점은 1852년 파리에서 개점한 ‘봉 마르세 (Bon Marche)’이며, 미국의 메이시(Macy, 1858년), 영국의 휘틀리(Whiteley, 1863년), 독일의 베르트하임(Wertheim, 1870년), 일본의 미쯔코시(三越, 1904년), 우리나라는 회신백화점(1931년)이 각 국의 최초 백화점으로 알려져있다(Park, 2015). Park(2015)에 따르면 초기 백화점은 소비자들이 도심 속의 거대한 건물에서 다양한 고품질의 상품들을 마음껏 돌아보면서 구매할 수 있도록 한다는 개념을 가지고 등장하였을 뿐 아니라 도시의 확장과 이로 인한 인구 집중으로 인해 점차 도시의 주요 소매업태로 정착되었다고 하였다(Park, 2015). 유통산업발전법에 백화점이라는 유통업태는 정의되지 않고 있으나 대규모점포 중 하나로 분류할 수 있으며, 백화점은 ‘부문별로 구성된 의류, 생활용품, 잡화, 식품, 가전 등의 각

종 상품을 매장 직원의 서비스를 통해 소비자에게 판매하는 매장면적 3,000㎡이상이고 직영비율이 50% 이상인 점포로 정의할 수 있다(Park, 2015).

우리나라 백화점의 발달과정은 <Appendix 1>과 같다.

우리나라 백화점은 코로나 19 로 인해 그동안 위축되었던 소비가 폭발하는 수혜를 입어 사상 유례없는 매출 호황을 누리고 있는 것이 사실이다. 그러나 최근 주력소비층뿐 아니라 팬데믹의 영향으로 익숙하게 된 비대면·언택트 소비성향의 확대와 이와 관련된 빠른 배송 등 다양한 편의시스템 제공은 향후 오프라인매장을 운영해야 하는 백화점들에게는 상당한 위험요인이 될 수 있다. 즉, 백화점이 온라인 친화적인 소비자들의 긍정적 경험(UPX: User Positive eXperience)을 오프라인 매장에서 어떻게 획득할 수 있는지에 대한 전략이 수립·운영되지 않는다면 백화점 또한 전통시장, 쇼핑몰 등 과거에 호황을 누렸던 업체들의 전철을 밟을 가능성이 매우 높다. 이런 시점에서 백화점의 문제점을 진단해보는 것은 매우 중요한 과정이 될 수 있다.

백화점 운영방식의 문제점으로는 먼저, 백화점의 직매입 비율이 너무 낮다는 점을 들 수 있다. 이는 재고 부담을 갖지 않는 장점도 있지만 백화점이 유통업태이기 보다는 점포 임대업이라고 분류해야 한다는 주장이 있을 정도로 수익실현을 위한 방향이 일반 유통업체와는 사뭇 다르다는 문제를 들 수 있다. 즉, 제조업체들이 백화점 입점을 선호하고 있기 때문에 시장과 소비자의 트렌드를 제대로 반영하지 않아도 일정 수익을 실현할 수 있음은 시장변화에 적극적으로 대응하는 유연성을 확보하지 못할 가능성이 있다는 것이다. 둘째, 특정매입매장과 임대매장에 적용되는 수수료율이 너무 높기 때문에 자기구매 현상이 빈번하게 발생하며 이런 문제는 언론에 빈번하게 보도되고 있는 것이 현실이다. 예를 들어 소수 백화점 업체에 의한 과점과 이로 인한 갑의 횡포로 인식되는 백화점에 대한 부정적인 이슈는 결국 백화점에게 부메랑을 작용하게 되어 백화점 브랜드에도 악영향을 미칠 가능성이 매우 높다. 또한 수수료율이 국내브랜드와 해외브랜드에 차별적으로 적용되고 있다는 점이나 입점업체가 수수료뿐 아니라 판촉비용이나 인테리어비용과 같은 부대비용을 부담한다는 점 등도 최근 이슈가 되고 있는 윤리경영이나 ESG (Environment, Social, Governance) 등 투명경영 철학을 위반하는 요인으로 작용될 수 있다.

한편 소비자에게 부담을 주는 문제점으로는 먼저, 도심지에 거액의 시설투자과 높은 경상운영비를 투입하기 때문에 책정되는 제품가격이 비싸다는 점을 들 수 있다. 또한, 과다한 판매비를 지출하는 적극적인 판매촉진 정책과 함께 과다한 판매촉진으로 인한 소비자들의 과소비를 부추는 등의 문제가 있다. 이런 이유는 우리나라 소비자는 아직까지도 백화점에서 판매하는

상품의 품질을 신뢰할 뿐 아니라 과시 소비와 같은 다양한 이유로 백화점 쇼핑을 선호하고 있기 때문이다.

본 연구에서는 그간 백화점에 대한 연구동향을 확인함으로써 위와 같은 문제점을 연구자들이 어느 정도 인식하고 있는지와 함께 문제해결을 위한 단서를 도출하고자 하였다.

2.2. BERTopic & LDA (Latent Dirichlet allocation)

자연어처리(NLP: Natural language processing)는 정보 관리, 의미론적 마이닝 및 컴퓨터가 텍스트 문서에서 인간 언어 처리로부터 의미를 얻을 수 있도록 하는 컴퓨터 과학의 도전적인 과제이다(Jelodar et al., 2019). 이 중 토픽모델링은 데이터 마이닝, 잠재 데이터 발견, 데이터와 텍스트 문서 간의 관계 찾기를 위한 텍스트마이닝에서 가장 강력한 기술 중 하나이다(Yang, 2022). NMF(Non-Negative Matrix Factorization)(Févotte and Idier, 2011)나 LDA(Blei et al., 2003)는 텍스트마이닝에서 다양한 장점이 있지만 최적의 결과를 얻기 위해 토픽의 수, 불용어 목록, 형태소 분석이 필요하고 단어 간의 의미관계가 무시된다는 약점을 가지고 있기에, 자연어처리 분야에서는 텍스트 임베딩(text embedding) 기술이 보급되고 있다(Yang, 2022).

2018 년 Google 이 공개한 BERT 는 입력데이터를 양방향에서 접근하여 문맥을 파악할 수 있고, 미세조정과 사전학습이 가능한 비지도 학습모델이다(Kim & Yang, 2022). BERTopic 은 이전의 복합 토픽 모델링(Bianchi et al., 2020; Devlin et al., 2018)과는 다르게 BERT 를 통해 사전학습된 언어모델 데이터를 호출해서 BoW(Bag of Word)의 요소를 제거하고 사전 학습된 언어모델을 적극적으로 활용하는 기법(Devlin et al., 2018)으로 최근 학계에서 주목을 받고 있다(Yang, 2022). BERTopic 은 상황에 맞는 단어 및 문장 벡터 표현을 생성하는 것으로 이러한 벡터 표현의 의미론적 속성은 유사한 텍스트가 벡터 공간에서 가까운 방식으로 텍스트의 의미를 인코딩할 수 있도록 한다(Grootendorst, 2022). 클러스터링 기술과 TF-IDF 의 클래스 기반 변형을 활용하여 일관된 주제 표현을 생성하는 BERTopic 은 텍스트 인베딩 단계에서 BERT 를 활용한 임베딩과 c-TF-IDF 단어 가중치를 활용한 다음, 각 도메인에 맞는 텍스트 클러스터링을 하여 텍스트에 잠재된 의미있는 주제를 찾아내는 토픽모델링 기법이다(Ko et al., 2022).

한편, LDA(Latent Dirichlet allocation)는 정형 데이터뿐만 아니라 비정형 데이터에서도 활용도가 높아 현재까지 토픽모델을 활용한 연구에서 가장 많이 활용되고 있는 토픽모델링 기법이다(Blei et al., 2003). LDA 기반의 토픽모델링은 자연어 처리, 텍스트마이닝 및 소셜미디어 분석과 정보 검색에 적용되었으며 이산 데이터를

시연하는데 탁월할 뿐 아니라 엄청난 정보에 숨겨진 구조와 의미를 찾기 위한 효과적인 접근을 제공하는 특성이 있다(Jelodar et al., 2019; Yang, 2022). LDA 는 말뭉치(corpus)의 생성확률 모델로 기본적인 아이디어는 문서가 잠재적인 토픽에 대해 무작위 혼합으로 포함되어 있으며, 토픽은 단어에 대한 분포로 구성된다(는 것이다(Jelodar et al., 2019). Blei et al. (2033)에 의해 처음 도입된 LDA 는 단어확률로 토픽을 나타내기 때문에 각 토픽에서 가장 높은 확률을 가진 단어는 일반적으로 LDA 토픽으로부터 단어확률이 무엇인지에 대한 좋은 아이디어를 제공하는 특징이 있다(Blei et al., 2013). LDA 매개변수(parameter)를 측정하기 위한 방법으로는 변이 방법(variational method)(Blei et al., 2003), 기대 전파(expectation propagation)와 Gibbs 샘플링(Griffiths & Steyvers, 2004)과 같은 방법들이 제안되었다.

본 연구에서는 BERTopic 이 다양한 장점이 있음에도 불구하고 임베딩 특성상 기존의 평가법으로는 BERTopic 의 성능을 평가하는 것이 불가능하다는 점(Bodrunova et al., 2020).에도 주목하고 있다. 실제 Yang(2022)과 Yang and Yang(2022)의 연구에서 보면 BERTopic 을 통해 도출한 토픽 수가 예상보다 훨씬 더 많아 이를 축소할 필요가 있었으며, BERTopic 에서 채택되지 못한 이상치 문서들이 다수 발생하여 이를 고려하지 않는 것은 실제 연구동향을 확인하고자 하는 목적에 부합하지 않다는 문제가 지적되었다. 따라서 본 연구에서는 Yang(2022)와 Yang and Yang(2022) 뿐 아니라 Ko et al. (2022), Kim and Yang(2022)이 적용한 것처럼 BERTopic 과 LDA 를 함께 활용하고자 하였다.

3. Research Procedure

백화점 연구동향 분석을 위한 연구절차는 <Figure 3>에 제시하였다.

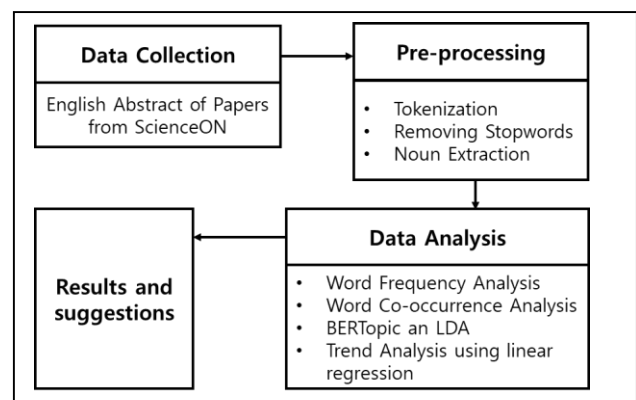


Figure 3: Research Procedure

3.1. Data Collection

2022년 10월 10일자 기준 scienceON (<https://scienceon.kisti.re.kr>)을 통해 키워드 "백화점"으로 검색한 결과 총 1,152 편의 논문이 도출되었으며 이 중 중복논문을 제외한 1,112 편 중 영문초록 551 편을 획득하였다. 연구자들의 발간 편수를 확인한 결과 성숙기로 구분한 2005~2009년, 2010~2014년 기간 동안 많은 연구들이 발표된 것을 알 수 있었다.

Table 1: Number of papers published on traditional markets by year

| Division | 1985~1999 | 2000~2004 | 2005~2009 | 2010~2014 | 2015~2019 | 2020~ |
|----------|-----------|-----------|-----------|-----------|-----------|-------|
| Num. | 35 | 100 | 147 | 136 | 102 | 31 |

3.2. Data Preprocessing

분석을 위해 데이터 전처리를 통해 단어 기준으로 토큰화하였다. 특수기호, 숫자와 함께 학술논문 초록에 빈번하게 등장하는 aim, research, purpose 등과 regression, coefficient, correlation 등 통계용어 등과 '백화점(department store)'을 불용어로 설정하고 이를 제거하였다.

4. Empirical analysis results

4.1. Results of Word Frequency Analysis

전체 데이터에 대한 단어 빈도분석 결과 'customer' (621), 'consumer' (543), 'service' (421), 'product' (411), 'space' (372) 등의 순인 것으로 확인되었다. 상위 20개 단어와 워드클라우드링 결과는 <Table 2>와 <Figure 4>에 제시하였다.

Table 2: Top 20 keywords as a result of word frequency analysis

| Rank | Keyword | Freq. |
|------|--------------|-------|
| 1 | customer | 621 |
| 2 | consumer | 543 |
| 3 | service | 421 |
| 4 | product | 411 |
| 5 | space | 372 |
| 6 | brand | 342 |
| 7 | satisfaction | 316 |
| 8 | market | 297 |
| 9 | image | 294 |
| 10 | sale | 287 |
| 11 | quality | 280 |
| 12 | design | 272 |
| 13 | shopping | 234 |
| 14 | relationship | 218 |
| 15 | value | 215 |
| 16 | type | 204 |

| | | |
|----|----------------|-----|
| 17 | area | 202 |
| 18 | system | 186 |
| 19 | company | 184 |
| 20 | characteristic | 182 |



Figure 4: Results of Word Clouding

4.2. Results of Word Co-occurrence Frequency Analysis

단어 간의 관계 확인을 위해 연결 중심성, 매개 중심성, 근접 중심성 및 고유벡터 중심성을 기반으로 단어 동시출현빈도를 분석한 결과 consumer, customer, marketing 등의 단어들이 product, service strategy, satisfaction 등의 단어들과 주로 연관되어 있는 것으로 나타났다. 동시출현빈도 상위 10개 단어는 <Appendix 2>에 제시하였으며, 연결빈도 40회 이상으로 단어 네트워크를 시각화한 결과는 <Figure 5>에 제시하였다.

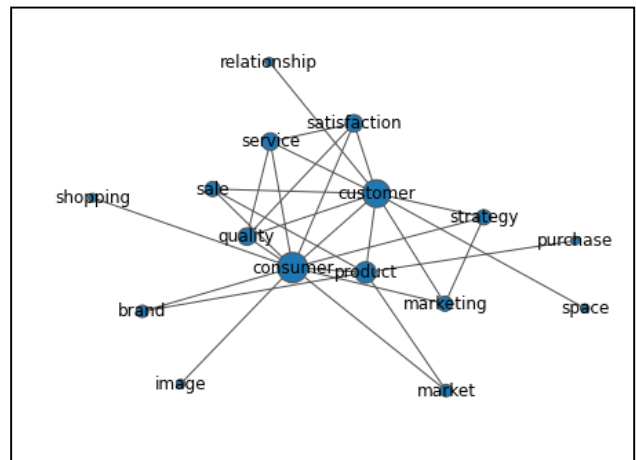


Figure 4: Results of Word Co-occurrence Network Visualization

<Figure 5>에서 보는 바와 같이 백화점과 관련된 연구들은 고객 또는 소비자를 중심으로 제품, 만족, 서비스, 품질, 마케팅전략과 같은 요인들에 대한 연구가 빈번하게 이루어진 것을 알 수 있었다.

4.3. Results of Topic Modeling

BERTopic 결과 토픽이 2 개로 분류되었으며 이상치 자료는 추출되지 않았다. 이런 현상은 그간 Yang(2022), Yang and Yang(2022)의 연구 결과와는 다소 다른 결과로 조심스러운 해석이 필요하다. 왜냐하면 BERTopic 은 상황에 맞는 단어 및 문장 벡터 표현을 생성하는 것으로 유사한 텍스트가 벡터 공간에서 가까운 방식으로 텍스트의 의미를 인코딩하기 때문에(Grootendorst, 2022), 유사하지 않은 텍스트는 이상치문서로 분류하기 때문이다. 분류된 첫 번째 토픽은 customer, consumer, service, product, brand 등 빈도분석에서 빈번하게 출몰한 단어들이며 총 538 편(97.6%)가 해당되었고, 두 번째 토픽은 observation, image, space, feature 등 점포 이미지와 관련된 단어들이 분류된 것으로 확인되었다. 따라서 본 연구에서는 토픽 1 을 LDA 를 통해 재분류하는 과정을 수행하여 토픽을 정교화할 수 있게 하였다.

<토픽 1>로 분류된 자료를 대상으로 LDA 를 실시한 결과 일관성(Coherence) 점수가 4 개 토픽일 때 0.3268 로 가장 높게 나타나 BERTopic 에서 1 개와 함께 총 5 개의 토픽을 확정하였다.

<토픽 1>은 observation, image, space, feature 등의 단어로 구성되어 점포의 외관과 이미지 등과 관련이 있을 것으로 판단하고 이를 '점포 이미지(shop image)'라고 명명하였다. <토픽 2>는 product, shopping, sale, information, system 등의 단어로 이루어져 판매와 관련된 쇼핑 정보와 관련이 있을 것으로 판단하고 이를 '쇼핑 정보(shopping information)'라고 명명하였다.

<토픽 3>은 service, satisfaction, quality 등의 단어로 이루어져 고객들이 백화점 서비스에서 인식하는 품질과 관련이 있을 것으로 판단하고 이를 '서비스 품질(service quality)'으로 명명하였다.

<토픽 4>는 brand, design, product, type 등의 단어로 이루어져 고객들이 인지하는 제품의 브랜드, 디자인 등 품질만족과 관련이 있을 것으로 판단하여 이를 '제품 만족(product satisfaction)'으로 명명하였다.

<토픽 5>는 market, sale, system, level 등의 단어로 이루어져 이를 '판매 시스템(sale system)'으로 명명하였다.

| | | |
|---|------------------------------|---|
| 2 | Shopping information (쇼핑 정보) | space, product, shopping, sale, consumer, customer, characteristic, good, information, system |
| 3 | Service quality (서비스 품질) | service, customer, satisfaction, quality, product, image, shopping, sale, area, relationship |
| 4 | Product satisfaction (제품 만족) | consumer, customer, brand, design, satisfaction, product, quality, value, type, service |
| 5 | Sale system (판매 시스템) | product, market, consumer, sale, customer, system, life, space, design, level |

4.4. Trend Analysis Results of Each Topic

토픽모델링을 통해 특정 논문에 할당된 각 토픽의 비율을 산출하고 산출된 비율과 각 논문의 발행연도를 활용하여 각 토픽의 연도별 평균비율을 계산하면 토픽의 트렌드를 확인할 수 있는데 이런 분석방법을 동적 토픽모델링이라고 한다(Yang, 2022; Yang & Yang, 2022). 동적 토픽모델링은 시간에 따라 토픽 비율이 유동적이기 때문에 각 토픽이 시간에 따라 등장하는 것을 추적할 수 있는 장점이 있다(Blei & Laffety, 2006; Yang, 2022). 동적 토픽모델링을 실시한 결과는 <Figure 5>에 시각화하였다.

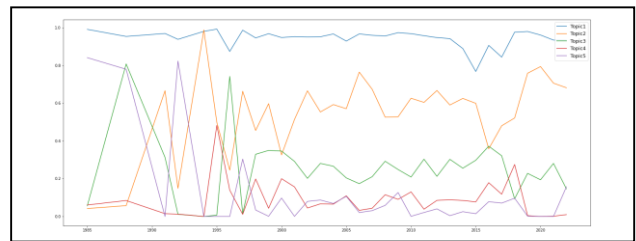


Figure 5: Visualization Results of Dynamic Topic Trend

<Figure 5>의 결과는 각 토픽이 연도별로 어느 정도 유동적인지는 확인할 수 있으나 연도별 증감에 대한 결과를 명확하게 판단하기에는 한계점이 있다. 따라서 본 연구에서는 Yang(2022)과 Yang and Yang(2022)이 적용한 분석방법을 활용하여 독립변수를 논문의 발행연도로 설정하고, 종속변수는 해당 연도 토픽의 평균비율로 설정한 후 OLS 회귀분석을 실시하였다(Yang, 2022; Yang & Yang, 2022). 회귀분석 결과가 통계적으로 유의미한 양(+)의 계수이면 관심도가 높은 토픽(Hot topic)으로, 음(-)의 계수이면 관심도가 낮아지는 토픽(Cold topic)으로 해석할 수 있으며, 통계적으로 유의하지 않는 경우에는 중립토픽(Neutral topic)으로 결론내릴 수 있다(Griffiths & Steyvers, 2004).

Table 2: Top 20 keywords as a result of word frequency analysis

| Topic | Topic Name | Keywords |
|-------|---------------------|--|
| 1 | Shop image (점포 이미지) | observation, image, space, feature, attention, intensity, time, section, characteristics, shop |

분석 결과 95% 신뢰도 구간에서 <토픽 2>의 '쇼핑 정보'(coeff. = 0.0097, $p < 0.01$)는 통계적으로 유의하게 관심도가 높아지는 토픽으로 나타났으며, <토픽 5>의 '판매 시스템'(coeff. = -0.0120, $p < 0.01$)은 상대적으로 관심도가 낮아지는 토픽인 것으로 확인되었다. 분석 결과는 <Table 3>에 제시하였다.

Table 3: Results of Regression Analysis

| Topic | Topic Name | Coefficient | t-value | p-value |
|-------|----------------------|-------------|---------|---------|
| 1 | Shop image | -0.0015 | -1.942 | 0.061 |
| 2 | Shopping information | 0.0097 | 3.079 | 0.004 |
| 3 | Service quality | -0.0012 | 0.403 | 0.689 |
| 4 | Product satisfaction | -0.0010 | -0.562 | 0.578 |
| 5 | Sale system | -0.0120 | -3.465 | 0.002 |

5. Conclusions

본 연구는 오프라인 유통업체에 대한 연속연구 중 하나로 백화점에 대한 연구동향을 파악함으로써 오프라인매장이 온라인쇼핑의 성장세에 대응할 수 있는 방안을 모색하는 것이다. 이를 위해 scienceON 에 등록된 논문의 영문초록을 Python 3.7 을 이용하여 BERTopic 과 LDA 를 통해 분석한 결과는 다음과 같다.

첫째, 단어 빈도분석 결과 customer, consumer, service 등이 빈번하게 출현하였으며, 단어 동시출현빈도를 확인한 결과 consumer, customer, marketing 등의 단어들이 product, service strategy, satisfaction 등의 단어들과 주로 연관되어 있는 것을 알 수 있었다. 이런 결과는 그동안 백화점과 관련된 연구들은 고객 또는 소비자를 대상으로 제품, 만족, 서비스, 품질, 마케팅전략과 같은 요인들에 대한 연구가 빈번하게 이루어진 것을 알 수 있었다. 둘째, BERTopic 결과 토픽이 2 개로 분류되어 이상치 자료가 추출되지 않아 첫 번째 토픽으로 도출된 538 편을 대상으로 LDA 를 실시한 결과 총 5 개의 토픽이 도출되어 이를 각각 '점포 이미지(shop image)', '쇼핑 정보(shopping information)', '서비스 품질(service quality)', '판매 시스템(sale system)'으로 명명하였다. 이런 결과는 그동안 많은 연구자들의 연구 결과가 특정 부분에 한정되어 있어 앞에서 지적한 백화점의 문제점을 해결하기 위한 정보를 제공하기에는 부족한 것이라 할 수 있다. 셋째, 각 토픽에 대한 트렌드분석 결과 통계적으로 유의하게 '쇼핑 정보(shopping information)'가 관심도가 높아지는 토픽이며, '판매 시스템(sale system)'은 상대적으로 관심도가 낮아지는 토픽인 것을 알 수 있었고, '점포 이미지(shop image)'의 경우 90% 신뢰도 수준에서 관심도가 낮아지는 것을 알 수 있었다. 백화점의 '판매 시스템'과 관련된 주제의 경우 3 사 과점체제인 현재 상황을 볼 때 백화점마다 큰 차이가 없는 것이 이유가 될 수 있을 것이며,

소비자들이 인지하는 백화점의 쇼핑과 관련된 정보에 대한 관심이 높은 것도 연구자들에게 긍정적인 영향을 미치는 것으로 해석할 수 있다. 그러나 이런 결과 또한 온라인매체와 경쟁을 벌여야 하는 백화점의 입장에서는 연구자들이 제공하는 정보가 다양한 단서들을 제공하는데는 한계점을 보이는 것이라 할 수 있다.

따라서 본 연구에서는 지금까지의 분석 결과를 토대로 향후 백화점을 대상으로 한 다음과 같은 연구주제에 관심을 가지기를 제안한다.

첫째, 유통업체로서 백화점의 진정한 역할에 대한 연구가 필요하다. 이는 백화점이 직매입의 낮은 비율과 함께 수수료를 기반으로 한 현재의 운영체제로는 유통업체로서의 긍정적인 역할을 기대할 수 없을 뿐 아니라, 높은 수수료율로 인해 부담을 갖는 입점업체나 높은 가격으로 인한 소비자들의 부담은 결국 백화점에게 부정적인 부메랑효과를 야기할 가능성이 높기 때문이다. 최근 '현대백화점 더 현대 서울'등이 체험과 경험을 중심으로 매장을 재구성하고, 명품브랜드의 입점보다 팝업스토어 등을 운영하는 등의 변화 시도는 소비자들의 백화점의 역할에 부응하기 위한 전략 중 하나라고 할 수 있다. 둘째, 백화점 차원에서 사용자의 긍정적 경험(UPX: User Positive eXperience)과 같이 소비자를 재해석할 수 있는 다양한 연구가 필요하다. 과거 백화점이 고급화를 통해 상류층이 이용하는 유통업체였다면 소비성향의 변화는 백화점이 대중적으로 접근할 수 있는 공간으로 해석되고 있는 것이 사실이다. 이런 이유 중 하나는 백화점 제품과 서비스에 대한 소비자들의 강한 신뢰가 기반이 된 것이라 판단된다. 즉, 온라인이나 타 유통업체에 비해 백화점 제품의 경우 믿을 수 있다는 소비자들의 신뢰로 인해 상당 분야의 제품들에 대한 경쟁력을 갖고 있다는 것이다. 따라서 소비자들이 인지하고 있는 백화점의 경쟁력과 상대적으로 부족한 부분에 대한 연구는 향후 백화점이 어떤 소비자들을 타겟(target)으로 설정하며, 제품이나 서비스 군을 어떻게 포지셔닝(positioning)해야 하는지에 대한 다양한 정보를 제공할 수 있을 것이다. 셋째, 백화점의 영업전략이 고객가치(customer value)와 고객평생가치(CLTV: Customer Life Time Value)를 기반으로 전환될 수 있는 연구가 필요하다. 백화점이 그간 지향해왔던 영업전략이 양질의 업체를 입점시켜 소비자에게 판매하도록 하고 이를 VIP 고객화하겠다는 전략이었다면, 이제는 다수의 방문 고객이 백화점을 자신만을 위한 공간으로 인식하게 하거나, 의미있는 체험을 통해 자신의 가치를 높일 수 있는 곳이라는 인식이 가능하도록 할 필요가 있다는 것이다. 즉, 방문고객의 평생가치를 기반으로 고객이 백화점을 방문했을 때 그곳만이 제공할 수 있는 가치 발굴을 위한

노력이 필요하다. 넷째, 백화점이 미래 유통 공간임을 소비자에게 각인시킬 수 있는 단서들을 찾는 연구가 필요하다. 이는 백화점뿐 아니라 오프라인 유통업체들의 공통적인 고민으로, 소비자들이 방문했을 때 항상 새로운 무엇인가에 대한 정보를 제공할 수 있어야 함을 의미한다. 따라서 다양한 업체와의 협업, 예체능 등 다양한 분야의 융합적 적용과 같은 새로운 신선함을 키워드로 한 전략적 접근은 소비자들을 고객화하고 긍정적 경험을 제공하는데 기여할 수 있다. 마지막으로 백화점이 윤리경영이나 ESG 등 투명경영의 효과성을 제고할 수 있는 다양한 주제들이 연구되어야 할 것이다. 높은 수수료율이나 판촉비용 등 부대비용을 입점업체에 관행적으로 부담하게 하는 백화점의 오래된 문제점을 고려해볼 때 투명한 지배구조(governance) 부분에 대해 취약할 가능성이 있으며, 이런 문제가 공론화되게 된다면 비윤리적인 기업으로 소비자들에게 외면당할 가능성 또한 높기 때문이다. 따라서 윤리경영을 기반으로 비재무적 요소로 최근 각광받고 있는 ESG 를 실천할 수 있는 다양한 사례연구와 실증연구들이 제시된다면 백화점이 소비자들을 고객화할 수 있는 전략 운영에 도움이 될 것이다.

본 연구가 오프라인 유통업체 중 백화점의 연구동향을 확인하고 발전을 위한 제안을 할 수 있었음에도 불구하고 몇 가지 한계점이 있어 향후 연구에서는 이를 보완할 필요가 있다. 첫째, BERTopic 의 결과 이상치문서가 추출되지 않고 2 개의 토픽으로 분류된 것들을 수 있다. 이런 결과는 Yang(2022)과 Yang & Yang(2022) 뿐 아니라 유통업체 분석을 위해 BERTopic 을 활용하는 연속연구 과정에서 처음 발생하는 것으로 BERTopic 의 성능을 평가하는 것이 불가능하다는 Bodrunova et al. (2020)의 지적과 관련이 있는 것으로 판단된다. 따라서 본 연구의 결과는 BERTopic 을 연구자들과 공유할 필요가 있으며, 전체 문서에 대해 LDA 와 BERTopic 을 통해 비교하는 연구 등을 통해 한계점이 있는지를 확인할 필요가 있다. 둘째, 토픽모델링이 갖고 있는 근본적인 문제인 토픽 명 결정의 한계를 들 수 있다. 비록 본 연구에서는 유통, 마케팅을 전공한 J 대학 교수들과 협업을 통해 토픽 명을 결정하였지만 토픽 명이 실제 연구되는 구체적인 내용을 모두 반영할 수 없다는 한계점이 있다. 따라서 향후 연구에서는 토픽으로 분류된 단어들을 최대한 아우를 수 있는 토픽 명 결정을 위한 방법에 대한 연구도 필요할 것이다.

References

- Bianchi, F., Terragni, S., Hovy, D., Nozza, D., & Fersini, E. (2020). Cross-lingual contextualized topic models with zero-shot learning. arXiv preprint arXiv:2004.07737.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan), 993-1022.
- Bodrunova, S. S., Orekhov, A. V., Blekanov, I. S., Lyudkevich, N. S., & Tarasov, N. A. (2020). Topic detection based on sentence embeddings and agglomerative clustering with markov moment. *Future Internet*, 12(9), 144. <https://doi.org/10.3390/fi12090144>
- Dennis, C., Murphy, J., Marsland, D., Cockett, T., & Patel, T. (2002). Measuring image: shopping centre case studies. *The International Review of Retail, Distribution and Consumer Research*, 12(4), 355-373.
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.
- Févotte, C., & Idier, J. (2011). Algorithms for nonnegative matrix factorization with the β -divergence. *Neural computation*, 23(9), 2421-2456.
- Gao, T. L., & Ko, H. J. (2017). A study on the last mile delivery for B2C fulfillment of fresh food e-commerce in China. *E-Trade Review*, 15(2), 51-71.
- Griffis, S. E., Rao, S., Goldsby, T. J., & Niranjani, T. T. (2012). The customer consequences of returns in online retailing: An empirical analysis. *Journal of Operations Management*, 30(4), 282-294.
- Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics, in Proceedings of the National Academy of Sciences, 101 (Supplement 1), 5228-5235. <https://doi.org/10.1073/pnas.0307752101>.
- Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv preprint arXiv:2203.05794.
- Groß, M. (2015). Exploring the acceptance of technology for mobile shopping: An empirical investigation among Smartphone users. *The International Review of Retail, Distribution and Consumer Research*, 25(3), 215-235.
- Holmes, A., Byrne, A., & Rowley, J. (2014). Mobile shopping behaviour: Insights into attitudes, shopping process involvement and location. *International Journal of Retail & Distribution Management*, 42(1), 25-39.
- Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent dirichlet allocation(LDA) and topic modeling: Models, applications, a survey. *Multimedia Tools and Applications*, 78(11), 15169-15211.
- Jiang, Y., Lai, P., Chang, C. H., Yuen, K. F., Li, S., & Wang, X. (2021). Sustainable management for fresh food e-commerce logistics services. *Sustainability*, 13(6), 3456; <https://doi.org/10.3390/su13063456>
- Kim, S. W., & Yang, K. D. (2022). Topic model augmentation and extension method using LDA and BERTopic. *Korean Society for Information Management*, 39(3), 99-132.
- Ko, Y. S., Lee, S. B., Cha, M. J., Kim, S. D., Lee, J. H., Ham, J. Y., & Song, M. (2022). Topic modeling insomnia social media corpus using BERTopic and building automatic deep learning classification model. *Korean Society for Information Management*, 39(2), 111-129.
- Kollmann, T., Kuckertz, A., & Kayser, I. (2012). Cannibalization or synergy? Consumers' channel selection in online-offline multichannel systems. *Journal of Retailing and Consumer Services*, 19(2), 186-194.
- Korcham (2021). *Distribution logistics statistics for 2021*. Seoul: Korea Chamber of Commerce and Industry Distribution and Logistics Agency.
- 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.
- MOTIE (2022). *Ministry of Trade, Industry and Energy. Annual '21, December '21 sales trends of major distribution companies*. Press release dated January 27, 2022.
- Nguyen, D. H., de Leeuw, S., Dullaert, W., & Foubert, B. P. (2019). What is the right delivery option for you? Consumer preferences for delivery attributes in online retailing. *Journal of Business Logistics*, 40(4), 299-321.

- Park, C. W. (2015). *Practical distribution theory*. Seoul: Cheongnam Book Publishing House.
- Shang, G., Pekgün, P., Ferguson, M., & Galbreth, M. (2017). How much do online consumers really value free product returns? Evidence from eBay. *Journal of Operations Management*, 53, 45-62.
- Tokar, T., Williams, B. D., & Fugate, B. S. (2020). I heart logistics—Just don't ask me to pay for it: Online shopper behavior in response to a delivery carrier upgrade and subsequent shipping charge increase. *Journal of Business Logistics*, 41(3), 182-205.
- Yang, H. C. (2022). Analysis of distribution industry research trends using BERTopic and LDA. *Journal of Creativity and Innovation (JCI)*, 15(4), 71-103.
- Yang, W. R., & Yang, H. C. (2022). Topic modeling analysis of social media marketing using BERTopic and LDA. *Journal of Industrial Distribution & Business*, 13(9), 39-52.
- Youn, M. K., & Kim, Y. O. (2017). *Distribution theory*. Seoul: Dunam Publishing House.

Appendixes

Appendix 1: Development process of Korean department stores

| Division | Period | Characteristics of growth process |
|---|--------------|--|
| Embryonic period (태동기) | 1930~1945 | <ul style="list-style-type: none"> - Japanese occupation(일제 강점기) - Opened Misskosi Gyeongseong branch (1930) and Georgia Gyeongseong branch (1931) (미스꼬시 경성지점(1930), 조지아 경성지점(1931) 개점) - Opened Korea's first Hwashin Department Store (1931) (국내 최초 화신백화점(1931년) 개점) |
| Period of confusion (혼란기) | 1946~1961 | <ul style="list-style-type: none"> - Absence of national commercial capital and the Korean War(민족 상업자본 부재와 한국전쟁) - Opened Midopa Department Store (1945), Dongwha Department Store (1955), etc. (미도파백화점(1945년), 동화백화점(1955년) 등 개점) |
| Period of dawn (여명기) | 1962~1972 | <ul style="list-style-type: none"> - Development of domestic industry and distribution system through the 1st and 2nd economic development plans (1,2차 경제개발계획으로 국내 산업 및 유통시스템 발전) - Direct operation of Shinsegae Department Store (acquisition of Donghwa Department Store) (신세계백화점 직영제 운영(동화백화점 인수)) - Acquired Midopa Department Store (Daenong Group)(미도파 백화점 인수(대농 그룹)) |
| Period of leap (도약기) | 1974~1979 | <ul style="list-style-type: none"> - Export-led high economic growth and distribution industry development policy (수출주도형 고도 경제성장과 유통업 육성 정책) - Lotte Department Store opens(롯데백화점 개점) - Increase in middle class consumers with purchasing power(구매력 갖춘 중산층 소비자 증가) - Vitalization of regional supermarkets (department store assistance) (지역기반 슈퍼마켓 활성화(백화점 보조)) |
| (Period of growing up) 성장기 | 1980~1995 | <ul style="list-style-type: none"> - Rapid economic growth due to the Asian Games, Olympics, etc. (아시안게임, 올림픽 등으로 인한 급속한 경제성장) - Convenience stores (1980s) and discount stores (1990s) appeared (편의점(1980년대), 할인점(1990년대) 등장) - Department stores expand in size, diversify and spread multi-store strategy (백화점의 대형화, 다각화 다점포 전략 확산) - Expansion of customer convenience facilities such as cultural centers (문화센터 등 고객편의시설 확충) |
| Transition period (전환기) | 1996~2004 | <ul style="list-style-type: none"> - Diversification and diversification of competitive businesses due to full opening of the distribution market (유통시장 전면 개방으로 인한 경쟁 비즈니스의 다각화·다원화) - Decreased consumption and high growth of hypermarkets due to IMF foreign exchange crisis (IMF 외환위기로 인한 소비 위축 및 대형마트 고성장) - Local department store closures and mergers and acquisitions between department stores (지방백화점 폐점 및 백화점 간 인수합병) - Implementation of department store operation efficiency and differentiation strategy (백화점 운영 효율화 및 차별화 전략 수행) - Establishment of specialty store and conversion to complex shopping mall(전문관 설립 및 복합쇼핑몰로 전환) |
| Period of maturity (성숙기) | 2005~2014 | <ul style="list-style-type: none"> - Strengthening promotional activities due to increase in foreign tourists (외국인 관광객 증가로 인한 판촉활동 강화) - Market stagnation due to late economic recession (후반 경기침체로 인한 시장 정체) - Accelerated entry of brands into the luxury goods market (명품시장 선정 관련 브랜드 입점 가속화) - CRM introduction and operation(CRM 도입 운영) - Intensifying competition with other retailers for key product groups (주력 상품군에 대한 타 소매업체와의 경쟁 심화) |
| (period of infinite competition) 무한 경쟁기 | 2014~present | <ul style="list-style-type: none"> - Improved presence of offline businesses such as convenience stores, discount stores, and complex shopping malls (편의점, 할인점, 복합쇼핑몰 등 오프라인 업체 입지 향상) - Expansion of online and mobile platforms, etc. (온라인·모바일 플랫폼 등 확대) - Changes in consumer lifestyle and consumption tendencies due to COVID-19 (코로나19로 인한 소비자 라이프스타일과 소비성향 변화) - Retailary consumption has recently spread after the consumption downturn caused by the pandemic. (팬데믹으로 인한 소비침체 후 최근 회복소비 확산) |

Appendix 2: Top 20 keywords as a results of Word Co-occurrence Analysis

| Rank | Word | Frequency | Rank | Word | Frequency |
|------|--------------|-----------|------|--------------|-----------|
| 1 | consumer | 72 | 6 | customer | 54 |
| | product | | | relationship | |
| 2 | customer | 68 | 7 | customer | 53 |
| | service | | | marketing | |
| 3 | marketing | 57 | 8 | consumer | 53 |
| | strategy | | | quality | |
| 4 | customer | 54 | 9 | quality | 51 |
| | satisfaction | | | service | |
| 5 | customer | 54 | 10 | satisfaction | 51 |
| | product | | | service | |