

# 텍스트 마이닝을 활용한 항공 화물 운임에 영향을 미치는 요인 분석: L.A. 및 프랑크푸르트 노선을 중심으로<sup>+</sup>

(Analysis of factors affecting air freight rates using text mining: focusing on the L.A. and Frankfurt routes)

최 동 현<sup>1)</sup>, 박 다 현<sup>2)</sup>  
(Donghyun Choi and DaHyeon Park)

**요 약** 본 연구는 2015년 1월부터 2021년 9월까지의 뉴스 기사를 분기별로 수집하여 토픽 모델링 분석 후 인천 출발 LA와 프랑크푸르트 노선의 항공 화물 운임과 회귀분석을 하여 항공 화물 운임 변동 예측에 영향을 미치는 요인을 분석하였다. 그 결과, 국내 물류 대비 LA 노선은 'Oil Price', 'Shipping' 토픽이 영향 있다는 점을 예측하였고, 프랑크푸르트 노선은 'Oil Price' 토픽이 영향을 미친다는 점을 결과로 제시하였다. 본 연구는 온라인 뉴스 기사 텍스트로부터 수집한 비정형 데이터를 활용하여 시장 참여자의 '보이지 않는 심리'에 영향을 주는 특정 키워드의 빈도 분석을 하고, 항공 화물 운임 사이의 동태적 관계를 파악하고자 하였다. 항공 화물 운임 예측은 항공사와 항공 산업 관련 다양한 분야에서 의사결정에 기초자료로 활용할 수 있을 것이다.

**핵심주제어** : 텍스트 마이닝, 토픽 모델링, LDA, 토픽, 항공 화물 운임

**Abstract** In this study, news articles from January 2015 to September 2021 were collected and analyzed to identify the topics that affect air freight rates through regression analysis based on freight rates and quarterly key topics for the Incheon to Los Angeles and Frankfurt routes. As a result, it was predicted that the "fuel" and "shipping" topics have an impact on the LA route compared to domestic logistics, and that the "fuel" topic has an impact on the Frankfurt route. This study aims the topics that influence the "invisible psychology" of market participants by analyzing the frequency of specific keywords in unstructured data collected from online news articles, and their relationship with air freight rates at specific times. Predicting air freight fares can serve as a foundational data for decision-making in various fields related to airlines and the aviation industry.

**Keywords:** Text Mining, Topic Modeling, LDA, Topics, Air Freight Rates

\* Corresponding Author: lily9373@naver.com

+ 본 논문은 석사학위 졸업논문을 바탕으로 작성되었음

Manuscript received March 23, 2023 / revised April 20,

2023 / accepted May 03, 2023

1) 중앙대학교 국제물류학과, 제1저자

2) 중앙대학교 무역물류학과, 교신저자

## 1. 서론

글로벌 경제성장과 교역환경 변화에 따른 국제 교역량 증가는 국제 교역의 주요 운송수단인 항공운송의 폭발적인 수요 증가와 항공운송산업의 지속적인 성장을 이끌어 왔다(Min and Ha, 2020). 이러한 가운데 코로나(COVID-19)로 인해 항공 산업은 큰 변화를 맞이하였고 항공 여객 수요는 급격한 감소가 되었으나, 항공 화물 수요는 공급 제한 속에서 반대로 성장하는 모습을 보였다. IATA 자료에 따르면, COVID-19의 장기화 속에서도 세계 총 화물 톤킬로미터(CTK)는 2019년 동월 대비 2021년도에 3.7% 증가하였다. 항공 화물은 현대사회의 많은 면에서 필수 불가결하다 할 수 있으며, 운송 상의 신속성이 필요한 신선식품(Perishable Goods)과 백신과 같은 의약품(Pharmaceutical) 등 콜드체인(Cold-Chain) 분야에서는 항공 운송에 대한 의존도가 높아지고 있다. 또한, 대부분의 사람들이 사용하고 있는 전자 제품들은 항공 운송과 연계되는 글로벌 공급사슬망(Global Supply Chain)에 의해 만들어지고 있으며, Amazon, Alibaba, e-Bay 등과 같은 E-commerce Web Sites는 항공을 이용한 특송 운송(Express)에 대한 의존도가 높아지고 있다(Lim, 2022).

이렇게 항공 화물 운송시장의 규모와 중요성이 확장되고 있는 상황에서 항공 화물 운임에 대한 예측이 가능하다면 항공 화물 운송 수단을 사용하고 있는 많은 Stakeholder들 중에서 특히 화주나 복합 운송업자들은 항공 화물 운임 변동 폭으로 인한 Risk를 조금이라도 줄일 수 있을 것이다(Kim, 2017). 항공 화물 운송 산업의 전망은 COVID-19 발생 이후에도 긍정적으로 평가되고 있으나, 항공 화물 수요를 결정짓는 주요 요인들의 영향력이 감소하고 팬데믹 관련 제한으로 인해 공급력에 압박이 가해지는 중이다. 항공 운송 산업은 다양한 외부 요인들에 의해 큰 영향 받기 때문에 항공 운송 산업의 특성상 항공 수요 예측은 매우 중요한 분야 중 하나이다.

COVID-19 발생 이후 항공 화물은 공간 공급의 제약 등으로 인해 운임 변동이 이상치를 보

여 시장 참여자들이 불확실성에 놓이는 상황이 되었다. 항공 화물에 대한 수요가 증가하면서 항공 운송 산업은 이에 대한 정확한 수요 예측의 필요성을 강조하고 있다. 항공 운송 산업의 수요 예측은 광범위하고 다양하게 연구가 진행되어 왔다. 그러나 대다수의 관련 연구들은 항공 화물보다는 여객에 치우쳐져 있다. 방법론으로는 전통적으로 시계열 모형을 활용한 연구가 주를 이루어 왔으나, 비정형 데이터인 텍스트 자료를 활용한 분석은 비교적 다른 분야에 비해 부족하다.

최근 몇 년간 다양한 분야에서 빅 데이터 분석을 통해 관련 주제를 연구하고 있다. 이러한 빅 데이터 분석을 활용한 대표적인 하나는 언론사의 뉴스 기사를 데이터로 활용한 연구이다. 언론사의 뉴스 기사는 특정 분야 및 정책 등을 분석할 때 대표적으로 사용되고 있는 원자료이고, 관련 분야 뉴스 기사들은 주요 내용을 독자들에게 전달함으로써 큰 영향을 미치기 때문이다(Kim and Kwon, 2020). 빅 데이터 분석 중 텍스트 마이닝 분석을 활용한 관광, 항공 분야의 연구에서는 비정형 데이터를 활용하여 그 속에 담긴 의미를 찾아내고 있다(Ko, 2021). 방법론 관점에서 빅 데이터에 대한 관심과 수요가 높아지면서 데이터 마이닝 기법들에 대한 연구가 증가하고 있는 추세이다.

본 연구는 뉴스 기사를 활용한 항공 화물 운임에 영향을 미치는 요인 분석 연구를 위해 텍스트 마이닝 연구 방법 중 토픽 모델링 LDA 기법과 통계적 분석 방법을 병행하였다.

본 논문의 구성은 다음과 같다. 2장에서 선행 연구를 검토하여 본 연구의 차별성을 제시하고, 3장에서 주요 연구 방법인 텍스트 마이닝과 토픽 모델링에 대해 설명한다. 4장에서는 연구 절차와 데이터 수집 방법에 대해 기술하며, 5장 변수 설명에 이어 6장에 분석 결과와 7장에서 결론을 제시하며 최종적으로 연구의 한계점 및 향후 과제까지 설명하고자 한다.

## 2. 이론적 배경

**2.1 항공수출 화물 물동량 현황(LA, 프랑크푸르트)**

우리나라 항공수출 화물 물동량은 상해, 홍콩, 베트남 노선이 L.A와 프랑크푸르트보다 많지만, 해당 지역들은 단거리 구간으로 장거리 구간에 비해 항공 화물 운임 변동 폭이 상대적으로 크지 않아 선행연구(Kim, 2017)를 참고하여 본 연구에서는 장거리 지역으로 범위를 한정하였다.

인천국제공항 출발 LA(LAX), 프랑크푸르트(FRA)의 2015년 1월~2021년 9월 기간의 항공수출 화물 물동량 현황은 Table 1과 같다.

LA(LAX) 노선의 경우 2016년 이후 기존 물동량보다 2018년도에는 약 15% 증가하는 것을 알 수 있으며, 지속적으로 증가하여 오히려 COVID-19 발생 이후 더욱더 항공수출이 활성화되어 2015년도에 비해 2021년도에는 80% 이상 증가한 것을 알 수 있다.

프랑크푸르트(FRA) 노선의 경우 2015년부터 변동은 있으나 COVID-19 이후 꾸준히 대략 10%씩 증가하고 있음을 확인할 수 있다.

Table 1 Air Export(LAX, FRA) Cargo Volume Status

Routes Period	Los Angeles (LAX)	Frankfurt (FRA)
2015	127,830	117,909
2016	122,439	116,138
2017	148,271	122,277
2018	148,301	121,447
2019	154,005	109,000
2020	183,608	129,201
2021	238,544	138,422

**2.2 항공 화물 수요 및 운임 예측 선행연구**

수요 예측은 항공운송산업의 특성상 중요한 분야 중 하나이다. 그렇기 때문에 항공 수요 예측에 대해 광범위하고 다양한 연구가 진행되어 왔다. 그러나 대다수의 항공 수요 예측 관련 연구들은 화물보다는 여객 분야에 치우쳐져 있으며, 방법론의 경우 시계열 모형을 활용한 연구

가 주를 이루어져 왔다.

다양한 수요 예측 선행연구를 연구 초점에 따라 1) 항공 화물 수요 예측 연구, 2) 항공 여객 수요 예측 연구, 3) 항공 운임 예측 연구로 나누어 설명하고자 한다.

첫 번째, 항공 화물 수요 예측 선행연구로는 Min and Ha(2020)은 데이터 마이닝 기법 중 의사결정 나무 방법론을 활용한 항공 화물 수요 예측을 하여 K 항공사의 일별 화물 물동량을 분석하여 수요를 예측하였다. Seo et al.(2017)에서는 시계열 분석을 이용하여 미주 항공 노선을 중심으로 항공 화물 수요를 예측하였다. Seo et al.(2009)에서는 회귀모형을 이용한 군항공 화물 수요 예측을 위해 물량과 요인들을 회귀분석을 이용한 연구를 하였다.

두 번째, 항공 여객 수요 예측 선행연구로는 Kim (2001)은 김포-김해 노선 항공 여객 수요를 예측하였을 때 2004년의 고속철도 개통이 영향을 미칠 것으로 예측하며 평균 성장 시나리오의 경우 김포-김해 여객 수요는 꾸준히 증가하는 것으로 결과를 제시하였으며, Kim et al. (2019)은 국내공항의 중·단기적 국제선 수요 예측을 위해 단 변량 시계열 모형인 ARIMA-Intervention 모형을 사용하여 2019년 국제선 항공여객 수요는 약 9.34% 증가할 것으로 항공 수요 예측 결과를 제시하였다.

세 번째, 항공 운임 예측 선행 연구에는 대표적으로 ARIMA(Autoregressive Integrated Moving Average) 연구 방법을 사용하고 있으나, Seo et al.(2014)에서는 기존 ARIMA 모형의 단점을 보완한 운임 예측 모형인 Error adjusted ARIMAC 모형을 활용하여 LA, Miami, Vienna의 3개 노선 항공 운임 데이터를 변수로 활용하여 미래 운임을 예측하였으며, Kim(2017)은 인천-미국 L.A 및 인천-프랑크푸르트 노선을 중심으로 항공 화물 수요에 영향을 미치는 요인을 분석하였으며, 해당 연구에서 독립변수로는 전월 항공 운임, 항공수출 물량, 환적 물량, 항공 유가, 성수기 여부를 사용하여 시계열 예측 방법으로 항공 화물 운임을 예측하였다. Ahn(2017)는 환율의 변화와 국내외 경기 수준이 해상 및 항공 운송에 미치는 동태적 영향

을 분석하였으며 해상 운송 수출입에 환율은 유의적 영향을 미치나 항공 운송에는 유의적인 영향을 주지 않는다는 것을 도출하였다.

기존 선행연구들은 항공 노선의 공급과 수요 중심 연구와 직접적으로 영향을 미치는 변수를 고려한 운임 예측 연구를 하였으나 수요와 공급 만으로는 설명하지 못하는 다양한 요인들이 운임에 영향을 미치고 있으며, 추가적으로 운임 및 가격 영향에는 시장 참여자의 심리가 반영되기도 한다.

그러하여, 본 연구에서는 기존 선행연구들과는 다르게 텍스트 마이닝 기법을 활용하여 항공 화물 운임에 영향을 미치는 요인을 연구해 보고자 한다.

기존 선행연구에서는 시계열 예측 방법론을 활용하여 항공 화물 운임을 예측한 연구가 주를 이루고 있지만, 최근 다양한 분야에서 연구되고 있는 텍스트 마이닝 연구 방법을 활용한 항공 화물 운임 분야에 대한 연구는 부족한 실상이므로 본 연구에서는 기존 연구의 한계점을 보완하고자 비정형 데이터를 활용한 연구를 해보고자 한다. 또한, 본 연구에서는 비정형 데이터를 활용하여 가격 분석 모형이 설명하지 못하는 부분을 분석하고자 한다. 키워드 분석을 통해 심리적인 요인들을 도출하여 운임에 영향을 미치는 요인들을 분석하고자 한다. 이 분석을 통해 요인별 영향력과 방향성을 예상하는 것은 항공 화물 시장 참여자들에게 매우 중요한 시사점을 제공할 수 있을 것이라 생각된다.

### 3. 연구 방법

#### 3.1 텍스트 마이닝(Text Mining)

텍스트 마이닝(Text Mining)이란 자연어로 구성된 비정형 텍스트 데이터에서 숨겨진 패턴이나 텍스트 간의 관계를 추출하여 활용 가치가 높은 정보를 찾아내는 의미 있는 분석 기법이다. 많은 양의 텍스트 데이터에서 의미 있는 정보를 추출하고, 연계성을 파악하며 텍스트의 카

테고리를 찾아내는 등 단순한 정보 검색 이상의 결과를 도출 가능한 기법으로 최근 많은 분야에서 텍스트 마이닝 연구 방법을 활용하고 있다(Kim et al., 2014). 텍스트 마이닝은 비정형 데이터 수집과 데이터 전처리 작업, 정보 추출, 정보 분석의 과정을 거치는 프로세스로 구성되어 있다.

#### 3.2 토픽 모델링 (Topic Modeling)

본 연구에서는 텍스트 마이닝의 프로세스 중 정보 분석을 위해 사용되는 기법 중 토픽 모델링 LDA(Latent Dirichlet Allocation) 기법을 사용하였다. 토픽 모델링(Topic Modeling)은 비정형 텍스트 자료에서 의미 있는 토픽을 도출해 내는 방법이다. 또한, 연관성이 높은 단어들을 그룹화 하여 각 그룹을 대표할 수 있는 토픽을 추론해 내는 방식이며(Park and Ahn, 2019), 특정 토픽의 중요도를 나타내는 통계적 수치를 통해 토픽별 비중도 확인할 수 있다(Choi et al., 2022).

토픽 모델링에는 다양한 방법이 있는데 LDA는 가장 널리 사용되는 것 중 하나이다. LDA 기반의 토픽 모델링 방법은 자연어 처리, 텍스트 마이닝 및 소셜 미디어 분석, 정보 검색에 적용된다(Blei et al., 2003). 토픽 모델링에서 토픽은 키워드의 조합이며, 각 키워드가 토픽에 일정한 가중치를 부여하게 되어 토픽의 키워드와 키워드의 가중치를 확인할 수 있다(Park and Ahn, 2019). 토픽 모델링에서의 토픽은 통계적으로 유의미한 방법에서 발생하는 단어 목록을 의미하며, LDA 기법의 대상은 이메일, 블로그 게시물, 저널, 기사 등 모든 종류의 구조화되지 않은 모든 텍스트가 될 수 있다.

LDA는 Corpus의 생성 확률 모델이며 LDA에 의해 도출된 토픽 모델들은 해당 텍스트 문서에서 단어의 의미와 개념을 이해하고 분석하는 것이 아니라 각 바구니(Baskets of words) 주제에 해당하는 단어를 선택하여 텍스트의 일부를 결합한다고 가정한다(Choi et al., 2022). LDA는 토픽을 담은 바구니에 가장 가능성 있게 분포될 때까지 이 프로세스를 계속 반복한다(Kim et

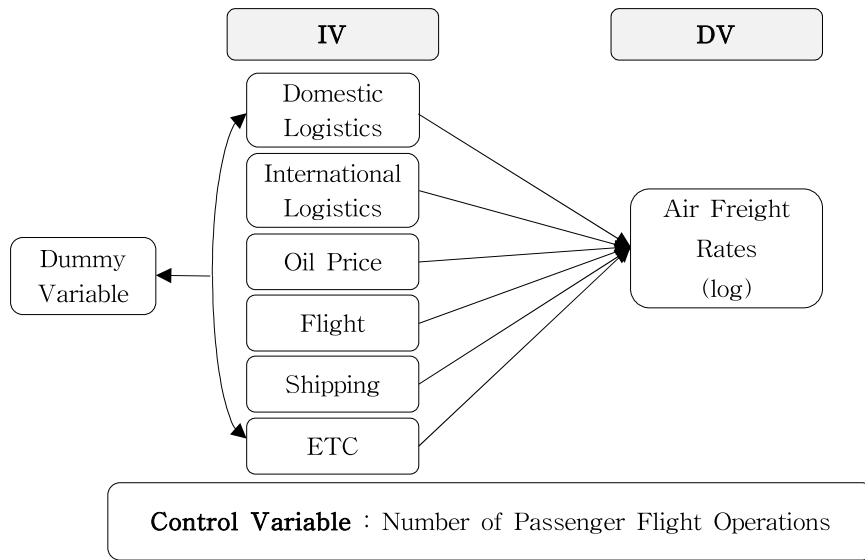


Fig. 1 Regression Research Model

al., 2003). 또한, LDA는 키워드의 확률로 토픽을 나타내며, 각 토픽에서 높은 확률을 가진 키워드는 일반적으로 해당 문서가 어떤 토픽인지 식별하는 데 좋은 아이디어를 제공할 수 있다 (Blei et al., 2003).

### 3.3 다중회귀분석

본 연구에서 사용하는 다중회귀분석의 경우 설명변수의 수가 많아질수록 변수의 유용성과 상관없이 기여율이 점점 높아지기 때문에 무의미한 변수에 대한 조정이 가능하도록 자유도로 조정된 기여율을 사용한다. 이것을 자유도 조정된 기여율 또는 Adjusted R<sup>2</sup>로 표기한다. 회귀식의 유의성을 보다 정확하게 판단하기 위해서는 무의미한 변수들을 판별할 수 있어야 하며, 이에 t 값과 p 값을 통해 유의성을 판단한다. 보통 t 값이 큰 변수일수록 목적 변수를 설명하는데 공헌도가 높다고 판단하며, t 값 대신 F 값을 이용하기도 한다. 또한, 유의 확률인 p 값을 통해 유의하지 않는 변수는 불필요한 변수라고 결론짓는다. 다중회귀분석에 있어서 중요한 점은 모든 변수가 필요하지와 불필요한 변수가 있는지를 검토하는 것이며, 회귀분석의 변수 척도가 등간척도나 비율척도가 아니더라도 더미

변수로 치환하게 되면 간단하게 회귀분석을 적용하여 영향력을 판단할 수 있다.

본 연구에서는 토픽 모델링을 통해 최종적으로 선정된 토픽들을 카테고리 별로 구분하여 더미 변수로 변환한 후 LA, 프랑크푸르트 항공 화물 운임과의 다중회귀분석을 진행하여 인천 출발 LA, 프랑크푸르트의 항공 화물 운임에 유의한 영향을 미치는 카테고리 간의 영향력을 파악하고자 한다. 이에 따른 회귀분석 연구 모형은 Fig. 1과 같다.

## 4. 연구 절차 및 데이터 수집

### 4.1 연구 절차

본 연구 분석을 위해 2015년 1월부터 2021년 9월까지 총 81개월 기간의 뉴스 기사 데이터를 한국언론진흥재단(Korea Press Foundation)에서 제공하는 빅 카인즈(Bigkinds) 뉴스 기사 분석 서비스를 활용하여 뉴스 기사 데이터를 수집하였으며, 기간별 핵심 키워드를 분석하고자 분기별로 나누어 뉴스 기사 데이터를 수집하였다.

특정 분야 및 정책 등을 분석할 때 대표적으로 사용되고 있는 언론사의 관련 분야 기사들은

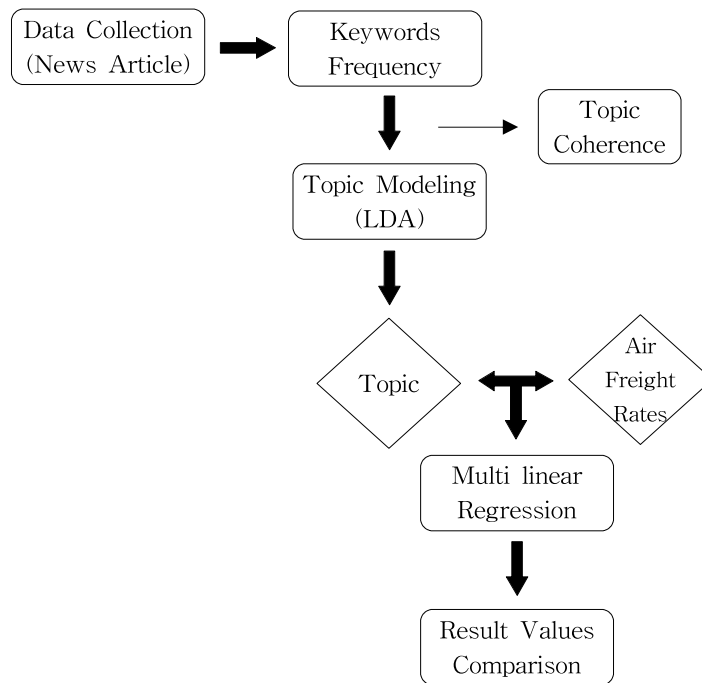


Fig. 2 Research Model

해당 시기의 주요 키워드들을 반영하기 때문에 다양한 연구 분야에 활용된다(Kim and Kwon, 2020). 국내 주요 일간지의 언론 보도를 수집하기에 앞서, 매체의 영향력을 고려하여 언론사를 선정하고자 하였다(Ham et al., 2021). 매체의 영향력은 국내 신문 구독률 순위(Korea Audit Bureau of Certification, 2022)를 기준으로 상위 6개 종합일간지를 선정하였다. 이에 선정된 언론사는 ‘조선일보’, ‘중앙일보’, ‘동아일보’, ‘한겨레신문’, ‘경향신문’, ‘한국경제’이다.

선행연구(Choi et al., 2012; Kim, 2017)에서 항공 화물 운임에 영향을 미치는 요인으로 제시한 자료를 참고하여 검색 키워드를 선정하였다. 선행연구를 바탕으로 본 연구에서는 ‘항공 화물’, ‘화물 운임’, ‘항공 수출’, ‘항공 수입’, ‘항공 Oil Price’, ‘물동량’ 총 6개의 중심 키워드를 선정하였다.

선정된 상위 6개 종합일간지에서 6개의 중심 키워드로 검색한 결과 뉴스 기사는 총 46,966건이었다. 검색된 뉴스 기사 자료에서 키워드를 추출하기 위하여 빅 카인즈에서 제공하는 형태소 분석 서비스를 이용하였다. 뉴스 기사 텍스

트 자료에서 키워드를 연구자가 직접 추출할 수 있으나, 다양한 언론사에서 작성한 기사들이기에 용어의 사용에 있어서 통일되어 있지 않으므로 선행연구(Na, 2020)에 따라 빅 카인즈에서 제공하는 형태소 분석을 통해 도출된 키워드를 사용하였으며, 중복 키워드를 가지고 있는 뉴스 기사를 제외하였다.

이후 빅 카인즈에서 제공하는 특성 추출 키워드를 이용하여 키워드 빈도수를 파악하였으며, 토픽 모델링 LDA 분석을 진행한 이후 회귀분석을 하였다. 이에 따른 연구 모형은 Fig. 2와 같다.

#### 4.2 뉴스 기사 데이터 수집

본 연구에서 주요 활용한 빅 카인즈(Bigkinds) 서비스는 중앙지 11개사, 경제지 8개사, 지역 종합지 28개사, 방송사 5개사, 전문지 2개사를 포함한 총 54개 언론사로부터 수집된 뉴스들로 구성된 통합 데이터베이스에 빅 데이터 분석 기술을 접목한 새로운 뉴스 분석 서비스이다. 빅 카인즈 서비스 시스템은 비정형 텍스트로 구성된 뉴스 정보를 정형화된 데이터로 전환하여 관련

사회현상을 다양하게 분석할 수 있는 기초 자료로 활용하는 데 도움을 주는 서비스를 제공하고 있다. 빅 카인즈 서비스의 구성은 뉴스 수집 시스템, 뉴스 분석 시스템, 뉴스 저장 시스템으로 구성되어 있다. 본 연구에서 주요 사용한 구체적인 연구 방법은 빅 카인즈의 뉴스 검색과 형태소, 개체명 분석과 빈도수 분석이다. 뉴스 기사 중 빅 카인즈에서 제공하는 특성 추출(가중치 순 상위 50개) 키워드를 이용하여 키워드 빈도수를 파악하였다.

### 4.3 항공 화물 운임 데이터 수집

본 연구에서는 인천 출발 LA 노선과 프랑크푸르트 2개 지역을 선정하여 운임 자료를 수집하였다. 우리나라 항공 화물 물동량 기준으로는 상해 및 홍콩, 베트남 지역이 2가지 노선 대비 물동량이 많지만 단거리 구간은 장거리 구간에 비해 항공 운임 변동 폭이 상대적으로 크지 않기에 선행 연구(Kim, 2017)를 바탕으로 미주 노선의 LA와 유럽 노선의 독일 프랑크푸르트로 설정하였다.

항공 운임은 성수기 및 비수기에 따라 계절성 운임 변동이 있기에 운임 예측 시 성수기 여부를 고려해야 하며, 결국 항공 운임은 거시경제 지표보다는 항공 화물 수요 및 공급에 직접적인 영향을 받는다고 할 수 있다(Kim, 2017). 본 연구에서 사용된 항공 화물 운임 데이터는 한국무역협회와, 한국관세물류협회에서 제공하는 항공 운임을 사용하였으며, 기타 부대비용을 제외하고 단순하게 항공 화물 운임만으로 규정하였다.

본 연구는 텍스트 마이닝 기법을 활용하여 국내 기준 뉴스 기사 분석을 통해 핵심 토픽을 분석하여 항공 화물 운임에 영향을 미치는 요인을 연구하는 데 의의를 두고 있으며, 이슈가 발생하고 즉각적으로 시장에 영향을 미치는 경우도 있겠으나 항공 화물 운임 자료는 분기별 흐름과 영향을 분석하고자 분기를 기준으로 정리하여 연구에 적용하였다.

## 5. 변수 설명

### 5.1 독립변수 및 종속변수

본 연구에서는 토픽 모델링에 이어 항공 화물 운임과의 회귀분석을 하고자 135개의 최종 토픽을 독립변수로 분류하기 위해 토픽 결과를 분석하여 추가로 6개의 카테고리 분류하였다. 카테고리는 전문가의 의견과 선행연구(Kim, 2017)의 연구 결과를 바탕으로 결정하였으며, 최종 항공 화물 운임에 영향을 미치는 변수로 ‘국내물류’, ‘국제물류’, ‘항공유가’, ‘항공’, ‘해운’, ‘기타’로 선정하였다.

‘국내물류’, ‘국제물류’는 항공 화물은 수요와 공급에 큰 영향을 받는다는 선행연구에 근거하여 세부적으로 분류하기 위해 국내와 International Logistics로 카테고리를 선정하였다. 또한, 항공 화물은 해상 운송의 대체제이기에 ‘Shipping’과 관련된 수요 공급에 따라 영향을 미치는 토픽이 다수였기에 ‘Shipping’ 카테고리로 분류하였다. 추가로 Oil Price는 항공 화물 운임에 가장 큰 영향을 미치는 요인 중 하나이며, 본 연구의 초점은 ‘항공’ 이기에 ‘항공’ 카테고리를 생성하였다. 그 외에 주식, 금리, 정치 등 항공 화물 운임과는 관련 없는 토픽들은 ‘기타’로 분류하였다.

본 연구에서는 종속변수로 사용되는 LA(LAX), 프랑크푸르트(FRA) 항공 화물 운임에 log를 취하여 사용하였으며, 독립변수는 토픽 카테고리를 더미 변수로 변환하여 사용하였다.

### 5.2 통제 변수 (항공 여객기 운항 편수)

통제 변수로 인천 출발 기준 LA(LAX), 프랑크푸르트(FRA) 노선의 여객기 운항 편수를 사용하였으며, 2015년 1월부터 2021년 9월까지의 여객기 운항 편수의 현황은 Table 2와 같다.

LA(LAX), 프랑크푸르트(FRA) 여객기 운항 편수를 비교하였을 때, 상대적으로 LA 노선의 여객기 운항 편수가 상당히 높다는 것을 확인할 수 있다. 또한, COVID-19가 발생한 2020년도 이후에 전체적으로 두 노선의 여객기 운항 편수가 감소하였으나 LA 노선에 비해 프랑크푸르트 노선의 여객기 운항 편수가 상대적으로 더 감소되었다는 것을 알 수 있다.

Table 2 Status of the number of flights operated by passenger aircraft

Routes Period	Incheon- L.A.	Incheon- Frankfurt
2015	8,248	1,324
2016	8,071	1,332
2017	8,493	1,328
2018	8,342	1,341
2019	8,531	1,384
2020	4,878	569
2021	3,527	371

## 6. 분석 결과

### 6.1 키워드 빈도 분석

2015년 1월~2021년 9월 기간에 ‘항공 화물’, ‘화물 운임’, ‘항공 수출’, ‘항공 수입’, ‘항공 Oil Price’, ‘물동량’을 동시 검색하여 뉴스 기사에서 많이 나타난 키워드들의 빈도를 분석하였다. 검색 키워드는 모두 ‘항공’이 공통적으로 포함되었기에 전체적인 기간의 상위 키워드는 ‘항공’이 나타났으며, 검색어가 모두 ‘항공’ 키워드가 포함된 점을 고려하여 각 기간에서 ‘항공’ 키워드는 제외하고 각 기간별로 뉴스 기사에서 많이 나타난 키워드 빈도를 다음과 같이 분석하였다.

2015년도에는 ‘중국’ 키워드가 170회로 가장 많이 나타났고, 다음으로 ‘한국’ 136회, ‘해운’ 112회 등의 순으로 나타났다. 2016년도에는 전년도와 다르게 가장 높은 빈도수의 키워드는 ‘해운’ 425회, ‘한진’ 186회, ‘한국’ 141회, ‘현대’ 139회 외에도 ‘한진 해운’ 키워드가 130회 등장하였으며 이는 2016년도에는 한진 해운 사태로 인하여 조선업의 침체, 구조조정 등의 이슈로 인하여 ‘해운’ 키워드가 가장 많은 빈도수를 차지하고 있음을 확인할 수 있었다. 2017년도에는 ‘해운’ 241회, ‘중국’ 127회, ‘인천’ 111회, ‘부산’ 110회 순으로 나타났다. ‘물동량’ 키워드가 중심 키워드이기에 ‘해운’과 관련된 키워드의 빈도수가 높으며 2017년도에는 부산항 사상 최초 물동량 2천만, 인천항은 300만의 TEU 시대가 열릴

정도로 전년 대비 물동량이 5.3% 이상 증가하였다. 이에 관련 키워드들의 빈도수가 높았다는 것을 알 수 있다.

2018년도에는 전년도와 유사한 키워드가 등장하였으며 2017년도와 동일하게 2018년도에도 해운과 항공의 화물 물동량이 증가하였으며 항공수출 화물 물동량도 꾸준히 증가하고 있다는 것을 확인할 수 있다. 2019년도에는 ‘공항’ 43회, ‘공사’ 42회, ‘물류’ 38회 등 항공 쪽에 더 가까운 키워드가 높은 빈도수를 차지하고 있는 것을 확인할 수 있다. 2020년도에는 ‘화물’ 68회, ‘해운’ 67회, ‘조선’ 59회, ‘물류’ 49회 등의 순으로 ‘물류’, ‘수출’ 키워드의 등장으로 2020년도에는 COVID-19의 영향으로 항공수출 화물 물동량이 COVID-19 발생 전인 2019년도에 비해 감소하였으나, FSC(Full Service Carrier)에 맞서 LCC(Low Cost Carrier)도 항공 화물 사업에 진출하기 시작하여 위와 같은 키워드의 빈도수가 높은 것으로 사료된다. 2021년도에는 ‘화물’ 53회, ‘물류’ 42회, ‘해운’ 41회, ‘택배’ 36회 등으로 등장하였다. 2020년도에 COVID-19의 직격탄을 맞은 항공사들의 항공 화물 운송량이 급감하였으나 LCC 항공사의 밸리 카고를 이용한 항공 화물 사업 진출로 인하여 잠시 주춤하였던 2020년에 비해 항공 화물 물동량이 회복된 것을 확인할 수 있다.

본 연구의 데이터 수집 과정에서 언론사를 구독률 순위 상위 6개 종합일간지를 선정하여 뉴스 기사를 수집하였기에 전체적으로 관련 뉴스 기사의 수에 대한 제한적인 부분이 있었다. 하지만, 광범위한 언론사 중에서 정확한 뉴스 기사를 추출하기 위해 구독률 순위를 적용하였다. 전체적으로 비슷한 패턴의 상위 키워드들이 등장하였으나, 연도별로 키워드 빈도 분석을 통해 뉴스 기사 트렌드 패턴을 유추할 수 있었다.

### 6.2 토픽 일관성 검증(Topic Coherence)

토픽 일관성 검증은 각 토픽들 안에서 높은 빈도로 도출된 단어들의 의미적인 일관성을 측정하는 척도이다(Stevens et al., 2012). 토픽 모델링을 평가하는 척도로는 혼란도(Perplexity)와



토픽 일관성(Topic Coherence)이 있다(Jo and Lee, 2021). 해당 모델이 얼마나 실제로 의미 있는 결과를 내는지 확인하고자 본 연구에서는 혼란도보다 더 이후에 제시된 방법론인 토픽 일관성을 사용하여 토픽 개수를 결정하고자 하였다.

토픽 일관성이 가장 높게 나온 지점이 토픽의 의미적 구분이 가장 잘 구현된 지점으로 볼 수 있으며, 그 지점을 최적의 토픽 개수로 판단하기도 한다(Stevens et al., 2012). 각 기간별 뉴스 기사 키워드 모음으로 Dictionary를 생성 후 Gensim의 Corpora를 이용하여 인코딩 후 Corpus를 생성하였고, Gensim의 Coherence Model을 활용하여

해당 Corpus의 최적의 Topic 수를 계산하였다.

2015년 1월~2021년 9월 기간의 연도별 토픽 일관성 검증 결과를 통한 최적의 토픽 지점은 5개로 선정하였으며, 검증 결과는 Table 3과 같다.

### 6.3 LDA 토픽 모델링 분석 결과

토픽의 개수를 Parameter로 LDA 토픽 모델링을 시행하였다. 최종적으로 분기별 토픽 일관성 검증 결과를 바탕으로 평균 5개로 산출하였으며, 각 토픽에는 10개씩의 키워드들이 구성되어 있으나 토픽 주제어 선정을 위하여 토픽 일관성 검증 결과를 바탕으로 중심적 의미를 가진 핵심 키워드 상위 5개를 선정하였다. 분기별 토픽을 추출하기 위하여 해당 기간의 뉴스 기사를 참고하여 분석하였으며, 추가로 항공과 해운 분야 전문가의 의견을 함께 고려하여 핵심적인 단어들 사이의 논리적인 연결을 통해 최종 토픽들을 선정하였다.

Table 4~10에서 2015년~2021년 기간의 LAX, FRA 운임의 변화 차이와 토픽을 한눈에 파악할 수 있으며, 매년 4분기에는 블랙프라이데이와 수요 및 공급의 증가로 인하여 항공 화물 운임이 모두 전 분기 대비 동일하게 상승하는 것을 확인할 수 있다. 본 연구에서 토픽 모델링 분석을 진행한 결과, 분기별 상위 5개 키워드는 총 675개이며 카테고리 별 토픽을 분기별(Q)로 분류한 결과 총 135개의 토픽들이 선정되었으며 결과는 Table 4~10와 같다.

Table 3 Topic Coherence Results

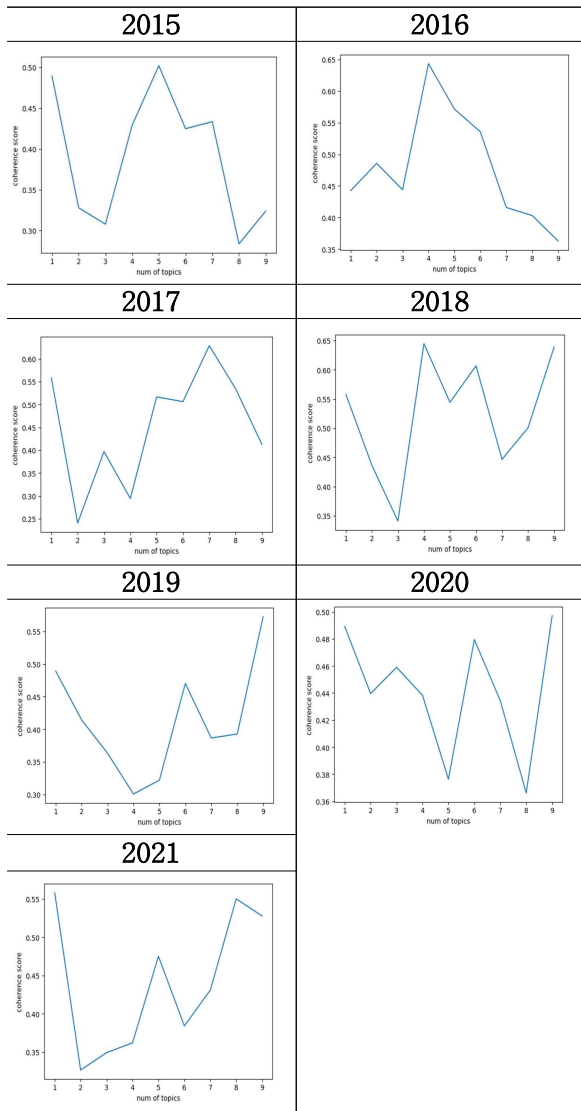


Table 4 Topic Categories by Quarter in2015

Q	LAX	FRA	Topic	Categories
1Q	3,900	3,067	Dependency on Logistics	International Logistics
			Cruise route development	Shipping
			Gapo New Port Development	Shipping
			Cambodia Southeast Asia Hub	International Logistics
			T'way Air	Flight

Analysis of factors affecting air freight rates using text mining: focusing on the L.A. and Frankfurt routes

2Q	4,200	2,700	Ara Kimpo Passenger Terminal	Shipping
			Parcel handling	International Logistics
			Asiana Airlines	Flight
			Daewoo Shipbuilding Demand Forecasting System Development	Shipping
			Low Oil Prices	Oil Price
			Marine Tourism	Shipping
3Q	2,767	2,700	Shipping Cargo	Shipping
			International Logistics linkage	International Logistics
			Second Suez Canal	Shipping
			Interest Rates and Economic Growth	ETC
			Gwangyang Port Investment	Shipping
			Shipping Cargo	Shipping
4Q	3,733	3,200	Ocean Freight	Shipping
			Cruise Tourism	Shipping
			Samsung Electronics	Domestic Logistics

Table 5 Topic Categories by Quarter in2016

Q	LAX	FRA	Topic	Categories
1Q	2,933	2,767	Automobile Transshipment port/hub	Domestic Logistics
			Eurasia	International Logistics
			Overseas Travel	Flight
			Low Oil Price	Oil Price
			Fuel Cost	Oil Price
			2Q	2,800
Gyeongbuk Logistics	Domestic Logistics			
Cargo ship	Shipping			
Busan Port Logistics Complex	Shipping			
Daewoo Shipbuilding (DSME)	Shipping			
3Q	2,900	2,533		
			Export	International Logistics
			Hanjin Shipping Crisis	Shipping
			Panama Canal	Shipping
			Low Oil Price	Oil Price
			4Q	4,067
Foreign Worker	ETC			
Dangjin Port	Shipping			
Aircraft	Flight			
New International Passenger Terminal	Shipping			

Table 6 Topic Categories by Quarter in2017

Q	LAX	FRA	Topic	Categories
1Q	3,600	3,100	Improper Solicitation and Graft Act	ETC
			Next-Generation Aircraft	Flight
			Musk Line	Shipping
			Translates to Incheon Airport Expressway	ETC
			Incheon International Airport	Flight
2Q	3,600	2,900	Logistics Cluster	International Logistics
			Lake Pyeongtaek	Shipping
			Local Tax Reduction	ETC
			Task Force	ETC
			Incheon Port	Shipping
3Q	3,500	2,800	Switzerland	International Logistics
			New Shipping Route	Shipping
			Integrated Logistics Center	Domestic Logistics
			Export	International Logistics
			Transportation Volume	Flight
4Q	4,533	3,300	Drone Delivery	Domestic Logistics
			Maritime Law	Shipping
			Smart Shipyard	Shipping
			Stock Price Decline	ETC
			Incheon New Port	Shipping

Table 7 Topic Categories by Quarter in2018

Q	LAX	FRA	Topic	Categories
1Q	3,967	3,200	Cheung Kong	ETC
			POSCO Daewoo	International Logistics
			Sulfuric Acid Cargo	International Logistics
			Summit Meeting	ETC
			Terminal 1	Flight
2Q	3,233	3,133	Hanjin	Domestic Logistics
			Parcel Rate	Domestic Logistics
			Oil Tanker	Oil Price
			Musk Line	Shipping
			Fuel Surcharge	Oil Price
3Q	3,600	3,067	Najin Port	Shipping
			Fuel Oil Spill	Oil Price
			LPG	International Logistics
			Incheon Port	Shipping
			Shipbuilding Industry	Shipping
4Q	4,367	3,667	West Coast Expressway	Domestic Logistics
			Interest Rate Hike	ETC
			Vietnam-linked Logistics	International Logistics
			Airlines	Flight
			Port Logistic	Shipping

Table 8 Topic Categories by Quarter in2019

Q	LAX	FRA	Topic	Cate gories
1Q	3,500	3,300	Incheon Airport Cargo Terminal	Flight
			Incheon Port	Shipping
			Manufacturing Business Industry	ETC
			New Airport	Flight
			Yantai Investment Briefing	ETC
2Q	2,700	3,333	Joint Venture	ETC
			Yeongilmanhang Station	Shipping
			LNG	International Logistics
			Hyundai Glovis	International Logistics
			Next-Generation Aircraft	Flight
3Q	2,833	3,667	Daiso Busan Hub Center	Domestic Logistics
			Mokpo New Port	Shipping
			Hormuz	Shipping
			T'way air	Flight
			Cargo Flight Increase	Flight
4Q	3,400	3,200	Stock Rally	ETC
			Incheon International Airport	Flight
			Export of Health Functional Foods	International Logistics
			Busan Port	Shipping
			Stock Rally	ETC

Table 9 Topic Categories by Quarter in2020

Q	LAX	FRA	Topic	Cate gories
1Q	6,200	3,612	Hong Kong Protests	Flight
			Turn black	ETC
			Exporting Batteries	International Logistics
			Shipping Industry	Shipping
			Busan Port	Shipping
2Q	10,198	4,837	Iron ore	International Logistics
			Lake Alhesiras	Shipping
			Airlines Industry	Flight
			Hyundai Mipo Dockyard	Shipping
			Hyundai Kia Motors	International Logistics
3Q	10,525	4,846	Sharp Drop in Fisheries Exports	International Logistics
			Incheon New Port	Shipping
			Block Chain	ETC
			Iron ore	Domestic Logistics
			Shanghai Container Freight Index	International Logistics
4Q	10,994	7,387	Delivery Industry	Domestic Logistics
			Black Friday	International Logistics
			Gadeok New Air Port	Flight
			Fulfillment	International Logistics
			Delta	Shipping

Table 10 Topic Categories by Quarter in2021

Q	LAX	FRA	Topic	Cate gories
1Q	11,134	8,217	Start Up	ETC
			Suez Canal	Shipping
			Baggage Claim Guide	Flight
			Gadeok New Air Port	Flight
			Incheon New Port	Shipping
2Q	11,213	7,625	Courier Strike	Domestic Logistics
			Export of Cosmetics	International Logistics
			International Balance of Payments	International Logistics
			Courier Strike	Domestic Logistics
			Suez Canal	International Logistics
3Q	11,579	7,874	Development of the West Sea	Domestic Logistics
			Shanghai Container Freight Index	International Logistics
			Increase in Domestic Flights	Flight
			Coupang	Domestic Logistics
			Market Insight	Domestic Logistics

6.4 회귀분석 결과

6.4.1 인천 출발 LA(LAX)

인천 출발 LA(LAX) 도착 항공 화물 운임과 토픽 간의 영향관계를 알아보기 위하여 다중회귀분석을 실시하였다. Reference Group 선정은 인천 출발이라는 점과 국내 신문사 기사에서만 토픽을 추출하였기에 “Domestic Logistics”로 선정하였다. 종속변수로 LAX 항공 화물 운임과 독립변수로는 International Logistics, ETC, Oil Price, Flight, Shipping을 분석하였으며, 통제변수로는 인천 출발 LA 운항 편수(여객기)를 사용하였다. 다중회귀분석을 이용하여 인천 출발 LAX 항공 화물 운임과 토픽의 회귀모형을 추정한 결과, 회귀모형은 통계적으로 유의하게 나타났으며(F=115.914, p<.001), 회귀모형의 설명력은 약 84.5%(R<sup>2</sup>=.845, adjusted R<sup>2</sup>= .837)로 회귀식의 설명력이 높게 나타났다. 분산팽창지수(VIF)도 모두 10미만으로 작게 나타나 다중공선성 문제는 없는 것으로 판단되었다. 회귀계수의 유의성 검증 결과, Oil Price(β=.089, p<.05), Shipping(β=.109, p<.05)이 LA 항공 화물 운임 변수에 대하여 정(+)의 영향을 미치는 것으로 나타났다. 추정결과 나타난 변수들의 회귀계수 의미는 다음과 같다.

Reference Group인 Domestic Logistics 대비 Oil Price 관련 토픽에 대한 검색량이 1단위(횟수) 증가하면 LAX 항공 화물 운임은 0.087(원)만큼 증가할 수 있는 것이라고 추정할 수 있으며, Shipping 관련 토픽에 대한 검색량이 1단위(횟수) 증가하면 LAX 항공 화물 운임은 0.050(원) 증가할 수 있다는 결과를 예측할 수 있다. 위의 결과는 선행연구(Kim, 2017)에서 연구한 결과와 일치하는 면이 있는데, 당월 항공유가가 LAX 항공 화물 운임에 영향을 미친다는 점이 연구 결과로 제시되었다. 이것은 항공 화물 운임과 Oil Price가 영향관계가 존재할 수 있다고 할 수 있다. 추가로, IATA분석에 따르면 항공사의 비용측면에서 유류비는 2017년 기준 항공사 총 비용의 22.4%로 가장 큰 비중을 차지하여 수익성과 밀접한 관계가 있다고 언급하였다(Korea Civil Aviation Association,

Table 11 LAX Regression Analysis Results

DV	IV	B	S.E.	$\beta$	t	p	VIF
LAX Freight Rates		8.935	0.209	-	42.773	0.000***	
	log (Number of LAX flight )	0.699	0.028	0.899	24.974	0.000***	1.067
	International Logistics	0.033	0.027	0.061	1.240	0.217	1.977
	ETC	0.040	0.029	0.064	1.365	0.175	1.794
	Oil Price	0.087	0.039	0.089	2.195	0.030*	1.345
	Flight	0.040	0.028	0.068	1.434	0.154	1.833
	Shipping	0.050	0.025	0.109	2.038	0.044*	2.358

F = 115.914(p<.001), R<sup>2</sup>= .845, adjR<sup>2</sup>= .837

Notes: a. \*p<.05, \*\*p<.01, \*\*\*p<.001

b. DV = LAX Air Freight Rates(log)

c. Reference Group = Domestic Logistics

2018). IATA는 아주 예외적인 컨테이너 운송 공급망 시장의 혼잡이 항공 화물 시장 수요에 영향을 주고 있는 것이 사실이라는 점을 언급하였다. 실제로 2021년 LA항만 물류적체로 인하여 항공 화물 수요가 상승하였다. 이 밖에 LAX 항공 화물 운임에 International Logistics, ETC, Flight 관련 토픽은 영향을 미치지 않는 것으로 다중회귀분석 결과가 나타났다.

이 점은 Domestic Logistics를 기준그룹으로 했을 때 Oil Price와 Shipping 토픽만큼의 유의한 영향을 미치지 않는다는 것으로 이해할 수 있다. 이에 따른 연구 결과는 Table 11과 같다.

#### 6.4.2 인천 출발 프랑크푸르트(FRA)

인천 출발 프랑크푸르트(FRA) 도착 항공 화물 운임의 Reference Group 또한 LA 항공 화물 운임 기준과 동일하게 비교하고자 “Domestic Logistics” 로 선정하였다. 종속변수로 FRA 항공 화물 운임과 독립변수로는 International Logistics, ETC, Flight, Oil Price, Shipping으로 하여 인천 출발 FRA 항공 화물 운임과 토픽 간의 영향관계를 알아보기 위하여 다중회귀분석을 진행하였다. 통제변수로는 인천 출발 유럽노

선 중 FRA 노선 운항 편수(여객기 기준)를 사용하였다.

회귀모형을 추정한 결과, 회귀모형은 통계적으로 유의하게 나타났으며(F=79,707, p<.001), 회귀모형의 설명력은 약 78.9%(R<sup>2</sup>=.789, adjusted R<sup>2</sup>= .779)로 회귀식의 설명력이 높게 나타났으나, LAX 노선에 비해 낮은 설명력이 나타났다. 분산팽창지수(VIF)도 모두 10 미만으로 작게 나타나 다중공선성 문제는 없는 것으로 판단되었다. 회귀계수의 유의성 검증 결과, Shipping( $\beta$ =.158, p<.0.5)이 FRA 항공 화물 운임 변수에 대하여 정(+의 영향을 미치는 것으로 나타났다.

추정 결과 나타난 변수들의 회귀계수 의미는 다음과 같다. Reference Group인 Domestic Logistics 대비 Shipping 관련 토픽에 대한 검색 량이 1단위(횟수) 증가하면 FRA 항공 화물 운임은 0.052(원) 증가할 수 있다는 결과를 예측할 수 있다. LAX 노선과 비교하였을 때, Shipping 토픽은 비슷하게 증가한다는 것을 알 수 있다. 이에 따른 연구 결과는 Table 12와 같다.

Kim(2017)에 의하면 LAX 노선과 함께 항공

Table 12 FRA Regression Analysis Results

DV	IV	B	S.E.	$\beta$	t	p	VIF
FRA Air Freight Rates		5.161	0.077	-	67.103	0.000***	
	log (Number of FRA flight)	0.282	0.014	0.864	20.495	0.000***	1.077
	International Logistics	0.035	0.022	0.091	1.588	0.115	1.977
	ETC	0.033	0.024	0.073	1.350	0.179	1.795
	Oil Price	0.057	0.033	0.082	1.734	0.085	1.353
	Flight	0.030	0.023	0.070	1.275	0.205	1.832
	Shipping	0.052	0.021	0.158	2.536	0.012*	2.361

F = 79.707(p<.001), R<sup>2</sup>= .789, adj R<sup>2</sup>= .779

**Notes:** a. \*p<.05, \*\*p<.01, \*\*\*p<.001

b. DV = FRA Air Freight Rates(log)

c. Reference Group = Domestic Logistics

화물 운임은 항공 유가에 큰 영향을 미친다고 하였으나, 본 연구는 국내 신문사에서 뉴스 기사를 수집하였기에 국내 한정된 뉴스 기사 자료에 대한 한계점이 있다는 점을 주장한다. 자세한 한계점에 대한 부분은 제 7장 결론 부분에서 제시하고자 한다.

## 7. 결론 및 시사점

본 연구는 빅 카인즈(Bigkinds) 뉴스 기사 데이터를 이용하여 항공 화물 시장 참여자들의 관심 키워드를 확인하여 인천 출발 LA, 프랑크푸르트 노선의 항공 화물 운임 변동 예측에 영향력을 미치는지 알아보려고 하는 것이 목적이며, 본 연구 목적에 부합하는 결과를 도출하였다.

본 연구의 주요 결론을 요약하면 다음과 같다. 2015년 1월부터 2021년 9월까지 81개월 기간 동안의 항공 수출 주요 지역 중 미국 LA와 유럽 독일 노선의 프랑크푸르트 항공 화물 운임 데이터를 종속변수로 선정하였다. 또한, 동일한 기간에 뉴스 기사를 제공하는 빅 카인즈(Bigkinds)를 활용하여 2021년 12월 31일 기준

국내 신문 구독률 상위 6개(조선일보, 중앙일보, 동아일보, 한겨레신문, 경향신문, 한국경제)를 선정하여 ‘항공 화물’, ‘화물 운임’, ‘항공수출’, ‘항공수입’, ‘항공유가’, ‘물동량’ 총 6개의 중심 키워드를 검색하여 관련 뉴스 기사를 추출하였다. 추출된 뉴스 기사를 바탕으로 빈도분석과 LDA 토픽모델링을 진행하여 핵심 토픽을 추출한 후 6개로 분류된 카테고리를 독립변수로 사용하였으며 추가로 여객기 운항 편수를 통제변수로 선정하였다. 토픽모델링 결과 카테고리 ‘Domestic Logistics’ 18개, ‘International Logistics’ 26개, ‘Shipping’ 44개, ‘Flight’ 21개, ‘Oil Price’ 7개, ‘ETC’ 19개로 총 135개이다. 독립변수로 사용되는 카테고리는 더미변수를 활용하여 항공 화물 운임과 회귀분석을 진행한 결과, 인천 출발 LA 노선의 회귀모형 설명력은 84.5%로 높은 설명력을 보이면서 인천 출발 LA 도착 항공 화물 운임은 Domestic Logistics 대비 ‘Oil Price’, ‘Shipping’ 토픽이 영향력을 미친다는 점을 결과로 제시하였다.

인천 출발 프랑크푸르트 노선의 회귀모형 설명력은 78.9%로 LA 노선 보다는 5.6%의 낮은 설명력을 보이지만 80%에 가까운 설명력을 나타냈다. 인천 출발 프랑크푸르트 도착 항공 화

물 운임은 Domestic Logistics 대비 ‘Shipping’ 토픽이 영향력을 미쳤다.

두 노선에서 ‘Shipping’ 토픽이 영향을 가장 미치는 점은 대륙에서 컨테이너로 화물을 한 번 운송하려면 평균적으로 Trailer transit time이 긴 상황에서 COVID-19으로 인해 Driver 수배가 어려워지고 컨테이너 회전율이 낮아지면서, 장비수급 부족 및 항만적체 심화되는 현상을 보이기 때문이다. 또한, COVID-19으로 인해 Vessel Schedule이 Skip되는 등 물류대란을 가중화시키며 이에 대한 대안으로 항공운송을 통해 긴급화물들이 운송되어진다. 이에 따른 수요 증가로 항공 화물 운임이 폭등하기 때문에 일정 시기에 ‘Oil Price’, ‘Shipping’과 관련된 토픽 검색 량으로 항공 수출 수요가 증가하는 것을 예측할 수 있다는 것이 본 연구에서 제시하고 있는 결론이다.

이러한 분석 결론들을 바탕으로 다음과 같은 학술적 및 실무적 시사점들을 제시하고자 한다.

첫째, 학술적 시사점으로는 기존 항공 화물 운임을 예측한 선행연구에서는 주로 시계열 분석 방법을 이용하여 항공 화물 운임 예측 연구가 진행되었으나, 본 연구에서는 항공 화물과 관련한 뉴스 기사 자료를 텍스트 마이닝 연구방법으로 항공 화물 운임연구에 적용하였다는 점이다. 또한, 기존 연구와 차별화하여 토픽을 이용한 항공 화물 운임에 영향을 미치는 요인을 파악한 후 추가로 회귀분석까지 진행하여 기존 선행연구와의 차별점을 가지고자 하였다.

둘째, 텍스트 마이닝 기법을 활용한 연구는 일부 기간의 핵심 이슈를 토픽으로 한눈에 파악할 수 있다는 이점을 가지고 있으므로 연구자들이 연구 방향성을 탐색하고 실무적으로 실무에 반영할 수 있는 토픽을 결정하는 데 활용될 수 있을 것으로 사료된다.

본 연구의 실무적 시사점으로는 비정형 데이터를 활용하여 시장 참여자의 보이지 않는 심리에 영향을 주는 매체 중 하나인 온라인 뉴스 기사를 활용하여 항공 화물 운임에 영향을 미치는 요인을 예측할 수 있으므로 항공 산업 관련 다양한 분야에서 의사결정에 기초자료로 활용할 수 있다는 점에서 의의가 있다.

## 8. 연구의 한계점 및 향후 과제

본 연구에서 연구 기간의 범위를 2015년부터 2021년 9월까지로 선정한 이유는 COVID-19으로 인하여 이후 운임의 변동 폭이 급증하여 2021년도 9월까지만 제공하였다는 점과 2022년 2월 러시아와 우크라이나 전쟁에 대한 여파로 인하여 항공 화물 운임이 급증한 특이한 상황으로 운임의 변동에 미치는 영향력이 정확히 파악이 되지 않을 것이라는 점을 고려하여 2022년도의 운임은 반영하지 않았다.

주식, 부동산 분야에서 해당 연구방법을 활용한 연구는 많으나 항공 화물 운임 분야에서 시장 참여자의 심리를 반영한 연구는 부족한 실상이다. 빅 카인즈(Bigkinds)에서는 굉장히 다양한 신문사의 뉴스 기사를 추출해낼 수 있으나, 본 연구에서는 광범위한 신문사를 모두 사용할 수 없었기에 국내 신문 구독률 상위 6개를 선정하였으며, 구독률 기준으로 선정함으로써 뉴스 기사가 제한될 수밖에 없었던 한계점이 있었다.

항공 화물 운임은 주식, 부동산 분야에서 주로 사용한 감성분석을 활용한 연구보다는 실제 상황을 제시하고 있는 뉴스 기사가 적합하다 판단이 되었으나, 최근 빅 데이터 연구의 증가로 감성분석 또한 잘 활용하면 추후에 더욱더 광범위한 연구를 할 수 있을 것이라 사료된다. 또한, 운임 기준은 인천 출발 LA와 프랑크푸르트 노선인데 뉴스 기사는 국내에서 제공하는 뉴스 기사들에서만 토픽을 추출했기 때문에 국내 중심이었다는 점이 한계점이며, 추후 연구에서는 물류 관련 사이트 혹은 구글 뉴스를 활용하여 전 세계적으로 노선에 영향을 주는 요인들을 파악하는 것이 더욱더 광범위하고 세분화된 연구를 진행할 수 있을 것이라 사료되기에 이는 후속 연구과제로 남겨두고자 한다.

## References

Airportal (<https://www.airportal.go.kr>)  
Ahn, K. A (2017). Analysis of the Effects of



- the Exchange Rate Volatility on Marine and Air Transportation. *Korea Trade Review*, 42(6), 131-154.
- Blei, D. M., A. Y. Ng and M. I. Jordan (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 993-1022.
- Cho, M. K. and Lee, B. J (2021). Comparison of service quality of full service carriers in Korea using topic modeling: based on reviews from TripAdvisor. *Journal of Hospitality and Tourism Studies*, 23(1), 152-165.
- Cho, C. H., S. H. Ahn, B. M. Park and S. M. Im (2012). Estimate and Forecast Air Freight Rates Using Stepwise Regression (From Incheon to LA and Frankfurt). *Korea Research Academy of Distribution and Management Review*, 15(2), 17-26.
- Choi, D. H., B. M. Song, D. H. Park and S. W. Lee (2022). Keyword trends analysis related to the aviation industry during the Covid-19 period using text mining. *Journal of the Korea Industrial Information Systems Research*, 27(2), 115-128.
- Choi, H (2021). Stock prediction analysis through artificial intelligence using big data. *Journal of the Korea Institute of Information and Communication Engineering*, 25(10), 1435-1440.
- Choi, J. H and Shin, C. S (2015). Correlation Analysis with Volume of Recreational Forest Visitors and Internet Search Words Using 'Big Data'. *Journal of the Korean Institute of Forest Recreation*, 19(4), 13-23.
- Ham, S. K., H. J. Kim and Y. W. Kim (2021). A Big-Data Analysis of Media Coverage on COVID-19 : Topic Modeling and Semantic Network Analyses by Issue Cycle and Political Orientation. *Korean Journal of Journalism & Communication Studies*, 65(1), 148-189.
- IATA (<https://www.iata.org/en>)
- Incheon International Airport ([www.airport.kr](http://www.airport.kr))
- Jeong, P. J., T. T. Zhao and H. S. Lee (2020). Analysis of Changes in the Air Passenger Transport Network by the Spread of COVID-19 -Focusing on International Airports in Asia-. *Korea Logistics Review*, 30(5), 119-136.
- Jeong, P. J., T. T. Zhao and H. S. Lee (2020). A study on the Change of International Air Cargo Networks at Asian Airports by the COVID-19 Pandemic. *Aviation Management Society of Korea*, 18(4), 71-87.
- Jin, H. G., U. G. Guk and K. W. Kang (2002). Air Cargo Demand Forecasting using Time Series Data. *Proceedings of the KOR-KST Conference*, 42(), 1-5.
- Korea Audit Bureau of Certification (<http://www.kabc.or.kr>)
- KITA ([www.kita.net](http://www.kita.net))
- Korea Customs Logistics Association ([www.kcla.kr](http://www.kcla.kr))
- Ko, M. H. (2021). A Study on Travel Bubble after Covid-19 Using Text Mining., *The Korea Academic Society of Tourism and Leisure Academic presentation proceedings*, 33-43.
- Kim, K. I. (2017). A Study on the Analysis and Prediction of Factors Determining Freight Rates for Korea's Airplane Export: Focusing on the L.A. Route in the United States and Frankfurt Route in Europe, *Inha University Logistics Graduate School Master's or Doctoral Thesis*.
- Kim, D. Y and Lee, Y. I (2018). News based Stock Market Sentiment Lexicon Acquisition Using Word2Vec, *The Korea Journal of BigData*, 3(1), 13-20.
- Kim, D. H and Cho, H. S (2020). An Empirical Study on Determinants of Freight Fluctuation in Air Transportation. *The Journal of shipping and logistics*, 36(2), 161-180.

- Kim, M. S (2001), A study on forecasting air passenger demand, *Journal of Aviation Development of Korea*, (2), 102-135.
- Kim, M. G., J. H. Ryu., D. H. Cha and M. K. Sim (2020). Stock Price Prediction Using Sentiment Analysis: from “Stock Discussion Room” in Naver, *The Journal of Society for e-Business Studies*, 25(4), 61-75.
- Kim, J. O and Kwon, C. H (2020). Comparative Analysis of News Articles related to Airlines and Staff the Previous Corona19(2019) and After Corona19(2020), *Journal of the Korea Society of Computer and Information*, 25(7), 167-173.
- Kim, J. C., H. G. Son and J. S. Park (2019). A Study on International Air Demand Forecasting by ARIMA-Intervention Model. *Journal of Korean Society of Transportation*, 37(1), 51-65.
- Kim, H. J., N. O. Jo and K. S. Shin (2015), Text Mining-Based Emerging Trend Analysis for the Aviation Industry. *Journal of Intelligence and Information Systems*, 21(1), 65-82.
- Kim, H. N., K. S. Lee and G. S. Jo (2003). Document Classification using Weighted Associative Classifier. *Korea Institute of Information Scientists and Engineers, Academic presentation proceedings*, 30(2), 154-163.
- Lee, M. C and Kim. H. J (2018). Construction of Event Networks from Large News Data Using Text Mining Techniques. *Journal of Intelligence and Information Systems*, 24(1), 183-203.
- Lee, J. Y and Ryu, J. P (2021). Prediction of Housing Price Index Using Artificial Neural Network. *Journal of the Korea Academia-Industrial cooperation Society*, 22(4), 228-234.
- Lim, J. P (2022). Analysis of air demand recovery and paradigm shift of the air transport industry after COVID-19. *International Journal of Tourism and Hospitality Research*, 36(2), 167-181.
- Min, K. C and Ha, H. K (2020). Forecasting the Daily Demand of Air Cargo Using Data Mining with CHAID Approach. *Journal of Korean Society of Transportation*, 38(3), 190-207.
- Moon, H. N and Kim, J. W (2014). A study on an individual stock return prediction model using internet news. *Korea Intelligent Information System Society Academic presentation proceedings*, 387 - 393.
- Park, G. Y and An, H. J (2019). The Topic Modeling Analysis of The DMZ Tour Issues Using Text Mining. *Journal of Tourism and Leisure Research*, 31(4), 143-159.
- Park, N. H and Kim, H. B (2016). The Effect of Keywords of Internet Search Engines on the Demand of Chinese Inbound Tourists: An Application of the Baidu Index Data. *Journal of Tourism Sciences*, 40(3), 159-174.
- Park, J. C., K. J. Han and H. Y. Chae (2019). Correlation Analysis between Livestock Mortality Caused by Heat Wave and News Big Data. *Journal of the Association of Korean Geographers*, 8(3), 529-543.
- Park, J. S and Lee, J. S (2021). Predictability of Housing Sales Prices Employing a Real Estate Sentiment Index : Using Unstructured Big Data of Online Newspaper and TV Broadcast News. *Journal of Korea Planning Association*, 56(4), 99-111.
- Rha, J. S (2022). Analysis of Factors Affecting Surge in Container Shipping Rates in the Era of Covid19 Using Text Analysis. *Journal of the Korea Industrial Information Systems Research*, 27(1), 111-123.
- Rha, J. S (2020). Analysis on Issues Related

to Supply Chain Management in the Era of Covid19 using Network Text Analysis. *Journal of the Korea Industrial Information Systems Research*, 25(6), 109-123.

Seo, S. K., B. H. Jung and I. K. Kim (2009). A Study on Estimating of Air Freight Demand using Regression Model. *Journal of the Military Operations Research Society of Korea (MORS-K)*, 35(3), 1-15.

Seo, S. S., J. W. Park., G. S. Song and S. G. Jo (2014). A Study of Air Freight Forecasting Using the ARIMA Model. *Journal of Distribution Science*, 12(2), 59-71.

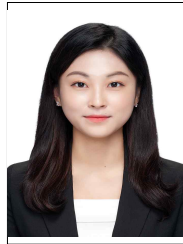
Stevens, K., P. Kegelmeyer, D. Andrzejewski and D. Buttler (2012). Exploring topic coherence over many models and many Paper presented at the 2012 joint conference on empirical methods in natural language processing and computational language learning of the *Association for Computational Linguistics*, 952-961.

Yoo, H. S and Kim, K. S (2021). A Study on Correlation Between COVID-19 and Postal Logistics Keyword Volume Using Machine Learning Techniques. *Proceedings of KIIT Conference*, 242-245



**최 동 현 (Donghyun Choi)**

- 정회원
- (현재) 중앙대학교 국제물류학과 부교수
- 관심분야: 혁신, SCM, 사회연결망



**박 다 현 (DaHyeon Park)**

- 정회원
- (현재) 중앙대학교 무역물류학과 박사과정
- 관심분야: 항공경영, 서비스 운영관리, 빅데이터, SCM