

빅데이터를 통한 2016년의 다이어트 실태 분석

정은진¹ · 장은재¹ · 조경애^{2,*}

¹동덕여자대학교 식품영양학과, ²고려대학교 보건과학대학

Analysis of dieting practices in 2016 using big data

Eun-Jin Jung¹, Un-Jae Chang¹, and Kyungae Jo^{2,*}

¹Department of Food & Nutrition, DongDuk Women's University

²College of Health Science, Korea University

Abstract The aim of this study was to analyze dieting practices and tendencies in 2016 using big data. The keywords related to diet were collected from the portal site Naver and analyzed through simple frequency, N-gram, keyword network, and analysis of seasonality. The results showed that exercise had the highest frequency in simple frequency analysis. However, diet menu appeared most frequently in N-gram analysis. In addition, analysis of seasonality showed that the interest of subjects in diet increased steadily from February to July and peaked in October 2016. The monthly frequency of the keyword highfat diet was highest in October, because that showed the 'Low Carbohydrate High Fat' TV program. Although diet showed a certain pattern on a yearly basis, the emergence of new trendy diets in mass media also affects the pattern of diet. Therefore, it is considered that continuous monitoring and analysis of diet is needed rather than periodic monitoring.

Keywords: big data, diet menu, exercise, commercial diet, continuous monitoring

서론

다이어트의 목적은 사회적 분위기와 시대적 인식에 따라 지속적으로 변화하고 있다. 최근에는 비만을 문제시 삼아 체중을 감량하고자하는 기본적인 목적에서 몸매를 관리하고 건강을 유지하기 위한 다이어트로 그 목적이 확대되었으며, 여성에게만 한정되었던 다이어트는 남성 그리고 청소년 등 다양한 계층으로 확산되었다(Van, 2007). 또한 다이어트는 사람뿐만 아니라 그 대상이 확대되고 있는데, Jung 등(2017)의 연구에 따르면 반려인을 중심으로 반려동물의 다이어트가 시도되고 있으며 그 경향은 사람과 비슷하다고 발표하였다. 이처럼 다이어트를 하는 대상이 다양해지고 있으며, 시기에 따라 각양각색의 다이어트 방법이 나타나고 있다.

이에 사람들은 여러 다이어트 정보를 얻기 위해 인터넷, 카페, 블로그 등 다양한 채널을 이용하고 있다(Lee, 2008; Oh, 2016). 특히, 인터넷은 대중매체 이용의 확장 및 연계의 공간으로서 자리매김하고 있고, 나아가 다양한 정보 교환의 수단으로 그 활용이 급증하고 있다(Cho와 Kim, 2018; Kim과 Lee, 2011). 따라서 다이어트에 관한 정보 흐름을 파악하기 위해서는 인터넷을 통한 데이터의 분석이 필요하다.

현재 데이터는 수치뿐만 아니라 문자, 영상 등 다양한 형태로 생성되고 있으며 생성 주기도 짧고 규모도 방대하여 대규모 데

이터를 의미하는 빅데이터를 형성하고 있다(Gordon, 2013). 이렇게 형성된 빅데이터는 다양한 분야에서 활용되고 있는데, 카드는 카드 소비 데이터를 가공하여 고객별 세그멘테이션 및 소비 경로 패턴 등에 활용하거나 결제 데이터를 통해 이상거래를 탐지하고 있으며(Chang, 2017; Jeong 등, 2015), 제조사는 생산성 향상과 품질관리 등에 데이터 분석 솔루션을 사용하고 있고, 통신사는 축적된 데이터를 활용해 상권 분석 및 타겟 마케팅을 지원하는 서비스를 제공하고 있다(Choi, 2016).

또한, 빅데이터를 통해 온라인이나 SNS에서 유통되는 글의 키워드나 검색 키워드의 분석 결과를 이용하여 다수의 사람들의 생각을 추측할 수 있다(Lee와 Hwang, 2014). 기존에는 사람들의 의견과 요구를 찾기 위해서 대부분 설문조사를 시행했다. 그러나 설문조사는 많은 비용과 시간이 소요되는 문제뿐만 아니라 제한된 표본으로 전체를 대변할 수 없는 문제점, 정해진 양식에 국한된 답변 등의 한계점이 존재한다(Seol과 Chung, 2000). 따라서 전체를 대변하고 자발적인 의견을 수집하여 분석할 수 있는 새로운 분석 방법인 빅데이터를 활용하고자 했다. 이를 활용하면 다이어트에 대한 사람들의 인식 및 변화의 흐름을 파악하여 다이어트 산업 분석, 다이어트 타겟 마케팅 등 다양한 분야에서 응용될 수 있다.

이에 Jung과 Chang(2016)은 빅데이터를 활용하여 2015년의 다이어트 현황을 다양하게 분석하였고, 그 뒤 Jung과 Chang(2018)은 2010년과 2015년의 빅데이터 비교분석을 통해 다이어트는 시대별 유행이 있음을 확인하였지만, 보다 정확하게 파악하기 위해서는 더 짧은 주기로 분석해야 한다고 언급하였다.

이와 같이 빠르게 변화하는 다이어트 유행과 사람들의 관심 등을 파악하기 위해서 2016년의 데이터를 분석하고자 한다. 이렇게 함으로서 기존의 논문들과 연관지어 1년 주기로 데이터의 연속

*Corresponding author: Kyungae Jo, College of Health Science, Korea University, Seoul 02841, Korea
Tel: +82-2-940-2764
E-mail: jo-kyung-ea@hanmail.net
Received February 15, 2019; revised March 15, 2019;
accepted March 15, 2019

적 비교가 가능하며, 또한 연속된 주기를 통해 다이어트에 관한 월별 결과를 파악할 수 있다.

따라서 본 연구는 포털 사이트인 네이버의 2016년 1월 1일부터 2016년 12월 31일 까지 1년 동안 ‘다이어트’ 키워드가 포함된 문장들을 수집했으며, 이를 통해 다이어트 관련 키워드 빈도 분석, N-gram 분석(다이어트 키워드와 동시 출현한 빈도와 방향성을 나타내는 분석), 키워드 네트워크 분석(유사한 단어끼리 군집을 형성하고 연관성을 파악하는 분석), 계절성 분석(다이어트 키워드 월별 출현빈도 분석)을 하여 2016년의 다이어트에 관해 실태를 분석하고자 한다.

재료 및 방법

연구 대상 및 기간

본 연구에서는 다이어트 현황에 대한 네트워크 분석을 실시하기 위해 국내에서 가장 인지도가 높은 인터넷 포털 사이트인 네이버를 분석대상으로 선정하였고, 데이터 수집을 위해서 키워드로 ‘다이어트’를 선정하였다. 다이어트는 시작하려는 날짜가 지정되어 있지 않으며, 개개인마다 다이어트를 시작하는 시기가 다르기 때문에 한 해가 시작되는 1월부터 끝나는 12월까지 분석하고자 했다. 이에 본 연구의 기간은 2016년 1월 1일부터 2016년 12월 31일까지 1년간을 설정하였다.

데이터 수집 및 전처리

데이터 수집을 위하여 Python 2.7 (Python Software Foundation, DE, USA) 프로그램을 이용하였다. 또한 네이버 검색 Application Programming Interface (API)의 기능을 이용하여 네이버 블로그, 웹문서, 뉴스 그리고 카페에서 ‘다이어트’ 키워드가 포함된 문장을 수집하였다. 수집된 문장은 한국어 특성에 맞게 최소한의 변경으로 형태소 분석이 가능한 온전한 프로젝트 Mecab (Unjeon, Korea) 프로그램을 이용하여 다이어트에 연관된 명사를 추출하였다.

데이터 분석을 실시하기 위해서, Mecab 프로그램을 통해 명사 형태로 분리된 키워드를 다음과 같은 전처리 작업을 통해 변환하였다. 첫째, 같이 사용되는 단어로 2개 이상 분리된 단어들은 다시 하나의 단어 형태로 변환하였다. 예를 들어 ‘닭가슴살’처럼 한 단어로 사용되는 단어가 ‘닭’, ‘가슴살’로 분리된 경우에는 다시 하나의 단어 형태인 ‘닭가슴살’로 변환하였다. 둘째, 단어의 형태가 약간 다르지만 동일한 뜻으로 사용되는 단어들의 경우에는 하나의 단어로 통일하였다. 예를 들어 ‘여성’, ‘여자’는 같은 뜻으로 간주하여 ‘여성’으로 통일하여 변환하였다. 마지막으로 ‘가’, ‘이’, ‘들’와 같이 조사 및 대명사 등으로 분리되어서 분석에 사용할 수 없는 형태는 삭제하였다. 이러한 작업을 통해 ‘다이어트’ 키워드를 포함한 연관키워드가 115,805개 나타났고, 총 단어의 빈도수는 13,926,052개가 도출되었다.

분석방법

수집된 방대한 양의 데이터 내용을 직관적으로 이해하기 위해서(Lee 등, 2014)는 주제 키워드와 연관된 키워드들의 출현횟수를 높은 순서대로 나열하여 분석하는 것이 유용하다. 왜냐하면 연관된 키워드들은 주제 키워드에 대한 관심과 흥미를 반영한 키워드로 간주할 수 있기 때문이다. 따라서 본 연구에서는 단순빈도분석을 통하여 수집된 데이터에서 출현 횟수가 높은 고빈도 키워드들이 무엇인지 분석하였다.

단순빈도분석을 통해서 데이터 전체의 빈도를 분석할 수 있으나, 주제키워드와 연관키워드 간의 동시 출현 및 밀집정도를 확

인할 수 없기 때문에 본 논문에서는 N-gram분석을 시행하였다. N-gram분석은 연구자가 지정한 n개의 어절 또는 음절을 연속적으로 분류한 후 단어들 사이에 밀집정도를 측정하고 빈도를 분석한다(Hwang 등, 2010). 예를 들어 N-gram을 통해 ‘다이어트 성공의 지름길은 식단조절이다.’라는 문장을 분석하고자 할 때 n=2이면 음절단위로 [‘다’, ‘이’], [‘이’, ‘어’], [‘어’, ‘트’], [‘트’, ‘ ’], [‘ ’, ‘성’], ..., [‘절’, ‘이’], [‘이’, ‘다’]로 분석한다. 만약 n=3이면 어절단위로 [‘다이어트’, ‘성공의’, ‘지름길은’], [‘성공의’, ‘지름길은’, ‘식단조절이다’], [‘지름길은’, ‘식단조절이다’, ‘ ’], [‘식단조절이다’, ‘ ’, ‘ ’]로 분석한다. 따라서 본 연구에서는 키워드단위로 분리하여 N-gram분석을 실시하였고, N-gram 네트워크 워드트리 를 이용하여 키워드 간 방향성도 확인하였다.

키워드 네트워크 분석은 문장에서 명사, 형용사 형태의 키워드들 간의 연관성을 파악하고 분석하여 키워드 간의 연결 관계를 연결망 형식으로 추출하는 방법이다(Diesner과 Carley, 2004; Ahn, 2012). 네트워크는 개체(actors)로 나타내는 노드(node)와 관계를 나타내는 링크로 구성되어 있으며(Moon, 2013), 본 논문에서는 키워드가 노드를 나타내고 키워드와 키워드의 연결이 링크로 나타난다. 이러한 키워드 네트워크 분석은 전체적인 네트워크 구조, 링크와 노드들 형태 및 특성, 노드들의 방향성 및 관계 등을 시각적으로 표현할 수 있다(Ahn과 Oh, 2015). 키워드 네트워크 분석을 위해 데이터를 소셜 매트릭스 형태로 구현해야 하는데 이는 (주)더 아이엠씨에서 제공하는 소셜 매트릭스 프로그램인 ‘텍스토크(The IMC, Daegu, Korea)’를 활용하였다. 또한 본 논문은 키워드 네트워크 분석 중 CONvergence of iterated CORrelations (CONCOR)분석을 이용하였다. CONCOR분석은 동시에 출현한 단어들을 매트릭스의 피어슨 상관관계에 따라 노드들의 블록을 식별하고 블록들 간의 관계를 파악하여(Wasserman과 Faust, 1994) 유사성을 지닌 키워들 간 관계를 이루고 군집을 형성하는 분석이며, 이를 시각적으로 나타내기 위해 Ucinet 6 (Analytic Technologies, Nicholasville, KY, USA) 프로그램을 통하여 시각화하였다.

다이어트에 대해 관심과 흥미가 많아지면 블로그와 카페 등 다이어트 키워드 자체의 노출 횟수가 많아질 것으로 간주하였고, 다이어트의 계절성 분석을 위해 월별로 다이어트 키워드의 출현 빈도를 분석하였다.

결과 및 고찰

단순빈도분석

다이어트 키워드와 관련된 키워드가 무엇인지 살펴보기 위하여 키워드 출현빈도를 알아보았다. 다이어트 키워드 검색 시 연관되어 출현한 키워드 중 상위 20위까지 선택하여 빈도표로 나타냈으며, 분석결과 운동이(191,032개) 가장 높은 출현빈도로 다이어트 키워드와 관련되어 있으며, 그 다음으로 식단(102,631개), 건강(86,199개)등의 순으로 나타났다(Table 1). Jung과 Chang(2016)의 2015년 결과에서도 ‘운동’ 키워드가 높게 나타났는데 그 이유는 문장 안에서의 ‘운동’ 키워드의 단순 반복이 높았다고 발표하였다. 이와 같은 이유로 본 논문에서 ‘운동’ 키워드가 높게 수집된 것으로 나타났다.

N-gram분석

키워드와 키워드 간 밀집정도를 통해 동시 출현빈도를 분석할 수 있는 N-gram 분석을 실시하였다. 그 결과, 다이어트-식단이 59,376개로 동시 출현빈도가 가장 높았으며, 다이어트-시작이 41,594개, 다이어트-성공이 37,964개 순으로 나타났다(Table 2).

Table 1. Frequency of 2016 keyword related to diet by simple frequency analysis

Rank	2016	
	Keyword	Frequency
1	Exercise	191,032
2	Menu	102,631
3	Health	86,199
4	Tea	85,602
5	Start	80,826
6	Effect	79,405
7	Breakfast	65,966
8	Today	62,117
9	Moment	60,327
10	Success	57,181
11	Diary	53,780
12	Management	50,154
13	Method	49,854
14	Kg	48,450
15	Farewell	47,735
16	Review	45,772
17	Dinner	45,452
18	Body	45,282
19	Thinking	43,200
20	Reduction	42,438

Table 2. Frequency of 2016 keyword related to diet by N-gram analysis

Rank	2016		
	Keyword1	Keyword2	Frequency
1	Diet	Menu	59,376
2	Diet	Start	41,594
3	Diet	Success	37,964
4	Diet	Exercise	33,363
5	Exercise	Diet	27,602
6	Health	Diet	24,082
7	Diet	Tea	23,261
8	Diet	Diary	22,996
9	Diet	Food	22,956
10	Diet	Method	22,345
11	Diet	Effect	21,485
12	Diet	Help	18,849
13	Diet	Moment	17,329
14	Diet	Clinic	15,956
15	Oriental medicine	Diet	15,279
16	Diet	Oriental medicine	15,207
17	Diet	Health	15,090
18	Today	Diet	13,360
19	Diet	Lunchbox	12,138
20	Start	Diet	11,890

Jung과 Chang(2016)과 비교해본 결과 2015년과 2016년의 N-gram 분석결과가 비슷하게 나타났다. 더군다나 상위 3위까지는 빈도만 다를 뿐 전년도와 동일한 다이어트-식단, 다이어트-시작, 다이어트-성공 키워드 순으로 나타났다. Lee와 Oh(2018)의 여대생들의 다이어트 인식에 대한 설문조사에서 다이어트 정보원으로 인터넷이 40%로 가장 높게 나타났으며, Kang과 Lee(2016)의 대학생의 인터넷 식생활 정보 이용실태를 알아보기 위한 설문조사에서도 다이어트가 가장 높은 관심을 보였다고 발표했다. 그리고 Kim 등(2016)의 연구에서 인터넷을 통한 식생활 정보로 다이어트 식단, 건강식단 활용 등을 이용하는 것으로 나타났다. 결국 다이어트를 시행하기 전 다른 것보다 식단을 우선 생각하고 그 후 다이어트의 시작과 다이어트의 성공을 생각한다고 볼 수 있다.

또한 Keyword1과 Keyword2를 이용하여 키워드 간 방향성을 알아볼 수 있다. Keyword1을 통해 다이어트라는 키워드가 먼저 언급되고 Keyword2를 통해 식단이라는 키워드가 언급되어 다이어트를 실시하기 위해 식단 키워드가 포함된 문장을 가장 많이 검색해봤을 가능성을 보여준다. 따라서 다이어트와 식단 키워드가 결합된 문장의 검색이 높았을 것으로 판단되었다. 또한 Jung과 Chang(2016)의 2015년에는 다이어트-식단(42,914개), 다이어트-차(28,917개), 다이어트-식품(17,564개), 다이어트-음식(11,949개) 등 음식에 관한 결과가 많이 나타났다. 이는 본 연구에서도 확인할 수 있는데, 다이어트-식단(59,376개), 다이어트-차(23,261개), 다이어트-식품(22,956개), 다이어트-도시락(12,138개) 등 해당 연구에서도 음식에 관한 결과가 많이 나타났다. Jang과 Kim(2016)의 연구에서는 최근 요리하는 방송인 ‘쿡방’과 먹는 방송인 ‘먹방’이 사람들의 집중을 받으며 음식이 화두가 되는 프로그램이 증가하고 있다고 발표하였다. Kang과 Lee(2016)도 식생활에 대한 방송 프로그램이 과거에 비해 많아지고 있고 결국 식단, 요리방법 등 식습관에 관한 관심이 높아지고 있다고 판단했다. 따라서 본 연구에서도 음식에 관한 결과가 높게 나타난 이유는 방송의 영향

을 받은 것으로 판단된다.

그러나 2016년에는 2015년과 다르게 다이어트-도시락(12,138개)이 나타난 것을 확인할 수 있다. Lee와 Cho(2016)에 따르면 도시락 시장의 규모는 2015년 2조 5천억 원에서 2016년 약 3조억 원을 전망하고 있다고 발표했다. Kim과 Sim(2017)의 네이버 트렌드 조사 결과에 따르면 다이어트 도시락 지표는 2011년에 비해 2016년 54배가 증가하여 다이어트 도시락에 대한 관심이 높아지고 있으며, 다이어트 목적, 1인 가구 증가, 온라인을 통한 편리한 배송 서비스 등이 도시락 소비를 증가시키는 것으로 파악했다. 또한, 기존의 도시락 형태와 다른 400칼로리 이하 다이어트 도시락, 채소 및 샐러드 도시락, 잡곡밥 도시락 등 건강과 다이어트를 추구하는 소비자를 위한 도시락이 출시되고 있는 추세이다(Kim 등, 2017). 따라서 본 연구에서도 2015년에는 나타나지 않았지만 2016년에는 다이어트-도시락의 결과가 나타난 것으로 판단된다.

키워드 네트워크 분석

CONCOR 분석을 이용하여 키워드 네트워크 분석을 확인하였다. 키워드 네트워크 분석은 총 4그룹으로 식이, 운동, 상업적 다이어트식품, 상업적 다이어트프로그램 그룹으로 분류되었다(Fig. 1). 식이 그룹으로 ‘식단’, ‘아침’, ‘점심’, ‘저녁’, ‘간식’, ‘쌀밥’ 등의 키워드가 묶여졌고, 운동 그룹으로 ‘운동’, ‘헬스’, ‘몸매’, ‘요가’, ‘건강’, ‘여름’, ‘추천’, ‘복싱’ 등의 키워드가 묶여졌다. 상업적 다이어트식품 그룹으로 ‘샐러드’, ‘닭가슴살’, ‘맛’, ‘차’, ‘가격’ 등으로 묶여졌고, 상업적 다이어트프로그램 그룹으로 ‘주비스’, ‘후기’, ‘의원’, ‘성공’, ‘관리’ 등의 키워드가 묶여졌다. Jung과 Chang(2016)의 2015년과 비교하면 그룹의 분류는 동일했지만 그룹에 속해있는 키워드에서 약간의 차이가 나타났다. 결과적으로 다이어트의 큰 틀은 식이, 운동, 상업적 다이어트 식품 그리고 상업적 다이어트 프로그램이 주를 이루고 있으며, 이 4개의 그룹에

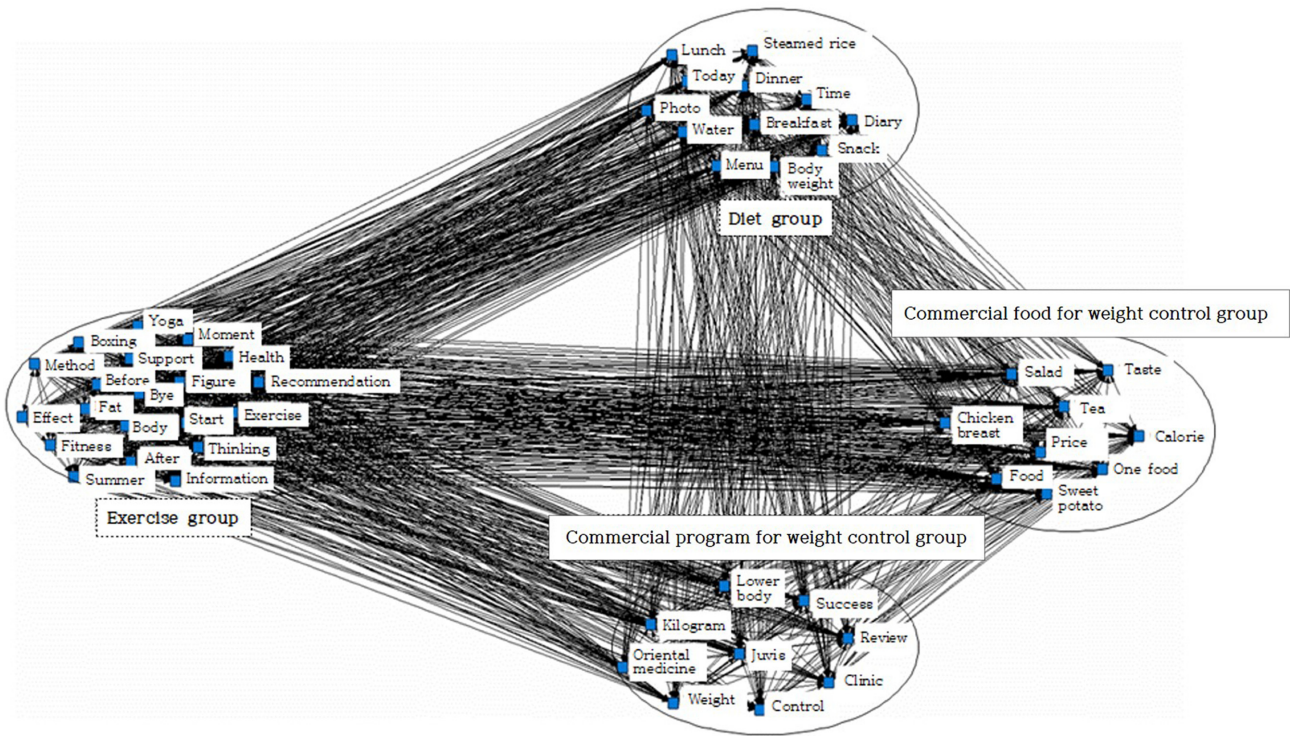


Fig. 1. Keyword network analysis related to diet in 2016



Fig. 2. Monthly frequency of keyword related to diet in 2016

속하지 않은 다른 방식이 존재하고 있다고 해도 그 존재가 아직까지는 미미하여 결과에 나타나지 않은 것이라 판단된다.

계절성분석

다이어트 키워드의 계절성 특징을 알아보기 위해서 출현빈도를 월별로 나누어 분석하였다. 분석결과 1월부터 12월 중 2월은 다이어트 키워드의 출현 빈도가 115,731개로 가장 낮게 나타났으며 출현 빈도는 7월까지 꾸준히 증가하여 7월은 142,482개로 높게 나타난 것을 확인할 수 있다. 그러나 1년 중 가장 높은 다이어트 키워드의 빈도수를 나타낸 달은 10월로 144,584개로 확인되었다(Fig. 2). 다이어트의 계절성 분석을 위해 월별 다이어트 출현빈도를 비교하였다. Jung과 Chang(2018)의 연구결과와 비교

해보면 2010년과 2015년 그리고 2016년의 그래프는 대체로 비슷한 양상을 보였는데, 이는 계절의 변화와 문화적 특성 때문인 것으로 판단된다. Pack(2004)과 Sul과 Park(2011)의 연구에서 겨울철에는 두꺼운 옷차림 때문에 몸매가 드러나지 않다가 여름이 다가올수록 옷차림이 가벼워지고 몸매가 드러나기 때문에 다이어트가 더욱 활발해진다고 발표하였다. 그리고 Kim(2016b)과 Lee(2016)의 연구에서 추석기간에 준비하는 음식은 강한 섭취 자극을 일으키며, 추석 음식 자체가 고칼로리 식품이기 때문에 다이어트의 동기 유발을 주장하였다. 이러한 선행연구와 동일하게 본 연구에서도 겨울이 끝나고 여름이 다가오는 2월부터 7월까지 꾸준히 상승하고, 추석연휴가 끝난 후 다시 출현빈도가 증가하는 것을 확인하였다. 그렇지만, 7월부터 10월까지 꾸준히 상승하지

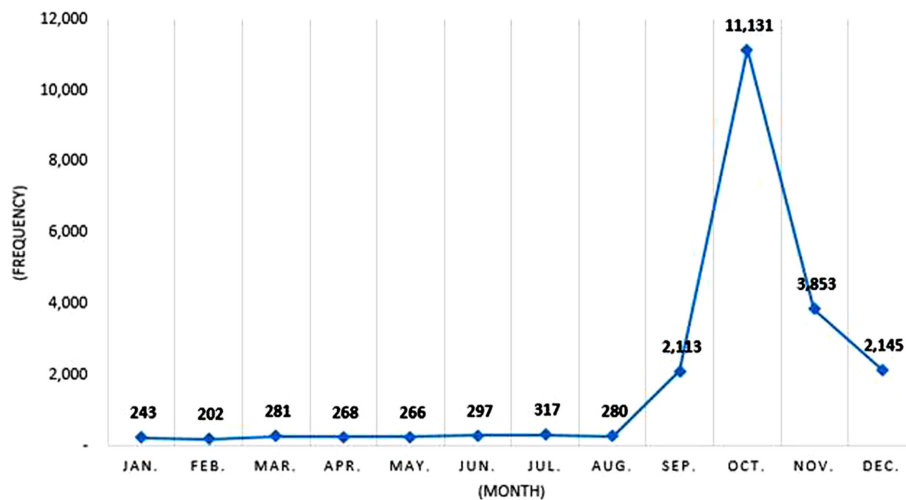


Fig. 3. Monthly frequency of high fat diet keyword in 2016

않는 것은 8월부터 여름이 서서히 끝나 다이어트의 관심이 낮아지고, 칼로리가 높은 추석 음식의 영향으로 10월에 다시 오르지만 이 관심 또한 오래가지 않아 11월부터 다시 낮아지는 것으로 판단된다.

마지막으로 Fig. 3은 9월 말 대중매체를 통해 화제가 된 고지방 다이어트의 출현빈도를 나타낸 그래프이다. 고지방 다이어트의 대중매체 방영 전인 1월부터 8월의 출현빈도는 미미하였으나, 방영 후 10월의 출현빈도는 11,131건으로 그 전보다 현저히 증가한 것을 확인 할 수 있다. 9월 말 문화방송(MBC)에서 ‘지방의 누명’이라는 다큐멘터리를 통해 고지방 저탄수화물 다이어트라는 대중들에게는 낯선 다이어트 방법이 소개되었는데, 고지방 저탄수화물 다이어트의 월별 출현빈도를 조사한 결과 10월에 급격한 상승이 나타나, 10월 다이어트 출현빈도 상승에 해당 프로그램이 영향을 주었을 것으로 판단된다. 고지방 저탄수화물은 케톤 식이요법에 근거한 다이어트 방식으로 탄수화물 섭취를 제한하고 버터, 치즈, 육류 등 지방을 마음껏 섭취한 후 이를 에너지원으로 사용하여 체중을 감량하는 방법인데(Kim, 2016a), 이 방법은 음식 섭취를 제한한 다이어트와 차별화 되어 섭취 제한 때문에 괴로워했던 사람들의 시선을 사로잡았다. Cho와 Kang(2017)과 Kang(2016)의 연구뿐만 아니라 다양한 매체에서 고지방 저탄수화물 다이어트에 관련된 프로그램 소개 후 사람들에게 저탄수화물 고지방 다이어트가 유행처럼 번지고 있다고 발표하였다. 따라서 대중매체 방영 후인 9월 말 이후부터 고지방 다이어트는 새로운 방식의 다이어트로 사람들에게 유행되었기 때문에 10월에 다이어트 출현빈도 수치가 고지방 다이어트의 출현빈도 수치가 높게 나온 것으로 판단된다.

그러나 영양 불균형, 요요 현상, 심혈관 계통 질환 유발 등의 부작용이 고지방 저탄수화물의 문제점으로 야기되었고, 결국 해당 다이어트 방법은 건강에 심각한 위험성을 초래할 수 있다고 한국영양학회를 비롯한 관련학회에서 공동 입장 표명이 있었다(The Korean Nutrition Society 등, 2016). 따라서 이 방법이 체중 감량에 도움이 될 수는 있어도 건강에는 악영향이 끼칠 수 있다는 인식으로 인해 그 유행은 오래 가지 못하고, 대중의 관심에서 금방 멀어졌으며, 본 연구의 그래프 또한 11월에 급격히 감소한 것을 확인할 수 있었다. 이처럼 이슈가 되는 다이어트들은 그 부작용과 역효과 등으로 인하여 대중들의 관심을 오래 유지하지 못하게 되며, 또 다른 방식의 새로운 다이어트가 등장하게 된다.

따라서 이를 모니터링하고 분석하기 위해 Jung과 Chang(2018)에서는 2010년과 2015년 5년의 주기로 분석하였지만, 분석주기가 너무 길어 그 사이에 이슈가 되었던 다이어트 방법들을 파악할 수 없었다. 특히, 다이어트는 계절뿐만 아니라 대중매체, 인터넷 등 다양한 경로를 통해 대중들에게 접근할 수 있기 때문에 1년 사이에도 큰 변화가 일어날 수밖에 없다. 그렇기 때문에 본 연구에서는 분석주기를 5년이 아닌 1년으로 줄여 분석하였고, 이 안에서도 월별로 큰 차이가 있는 것을 확인하였다.

결국 대중들에게 관심을 받고 이슈가 되는 다이어트 방법들은 빠르게 변화하고 있기 때문에 이를 파악하고 분석하기 위해서는 주기적이기 보다는 지속적인 모니터링과 분석이 필요하다고 판단되어진다.

요 약

인터넷과 대중매체의 발전은 새로운 다이어트에 대한 사람들의 접근을 용이하게 만들었다. 그러나 사람들의 관심은 시시각각으로 변화하기 때문에 이슈가 되는 다이어트는 매년 달라지고 있다. 따라서 본 연구에서는 2016년의 다이어트에 대한 경향을 알아보고 분석하기 위해서 빅데이터 분석 방법을 이용하였고, 포털 사이트 네이버를 통해 2016년 1월 1일부터 2016년 12월 31일 까지 1년간 다이어트 키워드가 포함된 문장을 수집하고 분석하여 단순빈도 분석, N-gram 분석, 키워드 네트워크 분석, 계절성 분석을 시행하였다. 단순빈도분석을 통해 가장 많이 출현한 키워드는 ‘운동’으로(191,032개)나타났고, 그 다음으로 ‘식단’이(102,631개)로 나타났으며, 키워드 간의 연관빈도를 분석한 N-gram 분석 결과 상위 결과로 다이어트-식단, 다이어트-시작, 다이어트-성공으로 나타났고, 다이어트-도시락이 새롭게 나타나 다이어트 시장의 새로운 변화를 확인할 수 있었다. 또한 다이어트 키워드와 연관된 키워드를 유사한 성격들끼리 그룹화한 키워드 네트워크 분석을 통해 식이그룹, 운동 그룹, 상업적 다이어트 식품, 상업적 다이어트 프로그램 그룹으로 총 4개의 그룹으로 세분화되었다. 계절성 분석을 통해 2월부터 7월까지 꾸준한 상승을 보였으나, 10월에 다이어트 출현빈도 수치가 급격히 상승하였고, 대중매체를 통해 소개된 고지방 다이어트의 월별 출현빈도도 10월에 급격한 상승이 있었다. 따라서 대중매체의 영향이나 새로운 다이어트의 유행이 사람들에게 큰 영향을 미치는 것을 확인할 수 있었

다. 이상의 결과를 바탕으로 다이어트의 패턴은 1년을 기준으로 일정한 양상을 띠고 있으나, 새롭게 유행하는 다이어트의 출현을 통해 사람들의 관심이 변화하여 다이어트의 패턴에도 영향을 미치는 것을 확인하였다. 결국 시시각각 변화하는 다이어트를 빠르게 파악하기 위해서는 주기적이기 보다는 지속적인 모니터링과 분석이 필요하다고 판단되어진다.

감사의 글

이 논문은 2018년도 동덕여자대학교 연구년제도에 의하여 수행된 것임.

References

- Ahn HJ. Extraction of keywords from brand images texts using network analysis. *JKIIT* 10: 176-182 (2012)
- Ahn MS, Oh IK. Analysis of attitudes on using five-star hotel packages applying network text analysis method. *InH. J. Tour. Manag. Sci.* 30: 163-181 (2015)
- Cho YG, Kang JH. Effectiveness and safety of low-carbohydrate diets. *J. Korean Med. Assoc.* 60: 40-46 (2017)
- Cho BK, Kim HJ. An empirical study on click patterns in information exploration. *J. DCS* 19: 301-307 (2018)
- Choi JK. Implications for domestic and overseas applications of big-data analysis. *KISTEP*. 14: 33-43 (2016)
- Diesner J, Carley KM. Causal mapping for research information technology. pp. 39-94. Idea Group Publishing, London, UK (2004)
- Gordon K. What is big data? *It Now* 55: 12-13 (2013)
- Hwang MG, Chop DJ, Lee HG, Choi C, Ko BK, Kim PK. Domain n-gram construction and its application. *JCSE* 37: 47-51 (2010)
- Jang YJ, Kim MR. Need for interaction or pursuit of information and entertainment?: The relationship among viewing motivation, presence, parasocial interaction, and satisfaction of eating and cooking broadcasts. *Korean J. Broadcast. Telecomm. Stud.* 30: 152-185 (2016)
- Jeong SH, Kim HN, Shin YS, Lee TJ, Kim HK. A survey of fraud detection research based on transaction analysis and data mining technique. *JKIISC* 25: 1525-1540 (2015)
- Jung EJ, Chang UJ. Tendency and network analysis of diet using big data. *J. Korean Diet Assoc.* 22: 310-319 (2016)
- Jung EJ, Chang UJ. Comparison and analysis of dieting practices using big data from 2010 and 2015. *Korean J. Community Nutr.* 23: 128-136 (2018)
- Jung EJ, Kim YS, Choi JW, Kang HW, Chang UJ. The analysis of a diet for the human being and the companion animal using big data in 2016. *Clin. Nutr. Res.* 6: 256-266 (2017)
- Kang SK. A low-carbohydrate, high-fat diet "fall into a new error". Available from: <http://weekly.donga.com/List/3/all/11/765463/1>. Accessed Oct. 21, 2016.
- Kang YE, Lee SY. Utilization of internet dietary information by university students in Seoul and Gyeonggi area. *Korean J. Human Ecol.* 25: 811-821 (2016)
- Kim JH. Effects of a low-carbohydrate, high-fat diet. *Korean J. Obes.* 25: 176-183 (2016a)
- Kim SY. Relationship between eating style and food intake of healthy female college students during chuseok holidays. *Korean J. Community Nutr.* 21: 131-139 (2016b)
- Kim JH, Kim YS, Jeon YJ. The effects of the convenience store lunch-box selection attributes on satisfaction and behavioral intentions. *JTLR*. 29: 353-370 (2017)
- Kim EM, Lee HM. Exploring audience behaviors in the two-screen environment. *Korean Assoc. Broadcast. Telecommunication Stud.* 25: 39-81 (2011)
- Kim NS, Lee KE, Son HJ. Comparison of health consciousness and smart-phone literacy according to the type of health information seeking behavior on the internet among one university students. *J. Health Info. Stat.* 41: 112-122 (2016)
- Kim B, Sim KH. Effects of food-related lifestyle on the importance of selected attributes of diet lunch box. *Korean J. Food Nutr.* 30: 413-426 (2017)
- The Korean Nutrition Society, Korean Society of Endocrinology, Korean Diabetes Association, Korean Society for the Study of Obesity, The Korean Society of Lipid and Atherosclerosis. Position of cooperation of five professional associations on the trend of eating of high-fat and low-carbohydrate. Available from: http://www.kns.or.kr/News/Notice_view.asp?mode=mod&resstring=%252FNews%252FNotice.asp%253Fsearch%253D1%253D%253Dcn_search%253D%25B0%25ED%25C1%25F6%25B9%25E6%253D%253Dxrow%253D10%253D%253Dpage%253D1&idx=670&page=1&xsearch=1&cn_search=%B0%ED%C1%F6%B9%E6. Accessed Oct. 26, 2016.
- Lee JH. Obtain diet information in online. *Medical Times*. Available from: <http://www.medicaltimes.com/Users4/News/newsView.html?ID=53684>. Accessed Jul. 27, 2008.
- Lee OR. High calorie chuseok food becomes extra fat?... chuseok diet method. *Munhwanews*. Available from: <http://www.munhwanews.com/news/articleView.html?idxno=23560>. Accessed Sep. 15, 2016.
- Lee DH, Cho WS. Another way of local food, Korean lunch pack. *RDA Interobang* 169: 1-20 (2016)
- Lee GH, Hwang SP. Business cycle indicator using big data: compilation of the naver search business index. *Economic Analysis* 20: 1-37 (2014)
- Lee SH, Oh BY. Concept mapping analysis of diet of college female students. *Korean J. Woman Psychol.* 23: 149-171 (2018)
- Lee OJ, Park SB, Chung DU, You ES. Movie box-office analysis using social big data. *J. Contests Assoc.* 14: 1-12 (2014)
- Moon JY. A study of the intellectual structure of secretarial studies using network analysis. *J. Secretarial Stud.* 22: 125-145 (2013)
- Oh HS. Body mass index and factors affecting on intention of dieting among college female students. *The Korean J. Advert. Public Relat.* 18: 137-167 (2016)
- Pack HJ. Various diet methods currently in vogue among the individual. *Korean J. Obes.* 8: 470-471 (2004)
- Seol HS, Chung TH. A study on methods estimating standard errors using replicated sampling. *J. Korean Educ.* 27: 98-111 (2000)
- Sul MS, Park DY. A prediction of demand for female sport participants by using seasonal ARIMA model. *JKPEAGW* 25: 179-192 (2011)
- Van JO. Food attitude and food behavior of female consumers according to their lifestyle. PhD thesis, Kyonggi University, Suwon, Korea (2007)
- Wasserman S, Faust K. *Social Network Analysis: Methods and Applications*. pp. 12-42. Cambridge university press, Cambridge, UK (1994)