The Analysis of a Diet for the Human Being and the Companion Animal using Big Data in 2016

Eun-Jin Jung,1 Young-Suk Kim,2 Jung-Wa Choi,3 Hye Won Kang,4 Un-Jae Chang1

1Department of Food and Nutrition, Dongduk Women's University, Seoul 02748, Korea
2Department of Food and Nutrition, Ansan University, Ansan 15328, Korea
3Department of Food and Nutrition, Soongeu Women's College, Seoul 04628, Korea
4Department of Family and Consumer Sciences, North Carolina A&T State University, Greensboro, NC 27411, United States

ABSTRACT

The purpose of this study was to investigate the diet tendencies of human and companion animals using big data analysis. The keyword data of human diet and companion animals’ diet were collected from the portal site Naver from January 1, 2016 until December 31, 2016 and collected data were analyzed by simple frequency analysis, N-gram analysis, keyword network analysis and seasonality analysis. In terms of human, the word exercise had the highest frequency through simple frequency analysis, whereas diet menu most frequently appeared in the N-gram analysis. Companion animals, the term dog had the highest frequency in simple frequency analysis, whereas diet method was most frequent through N-gram analysis. Keyword network analysis for human indicated 4 groups: diet group, exercise group, commercial diet food group, and commercial diet program group. However, the keyword network analysis for companion animals indicated 3 groups: diet group, exercise group, and professional medical help group. The analysis of seasonality showed that the interest in diet for both human and companion animals increased steadily since February of 2016 and reached its peak in July. In conclusion, diets of human and companion animals showed similar tendencies, particularly having higher preference for dietary control over other methods. The diets of companion animals are determined by the choice of their owners as effective diet method for owners are usually applied to the companion animals. Therefore, it is necessary to have empirical demonstration of whether correlation of obesity between human being and the companion animals exist.

Keywords: Data analysis; Diet; Pets; Dogs; Cats; Humans

INTRODUCTION

As health consciousness is on the rise around the world, new diet trends emerge and old diet trends get disappeared. In the 1980s, one-food diet, consuming only a single type of food or fasting, was in trend. Enzyme diet using vegetables, and Atkins diet composed of high-protein and low-carbohydrate food, were popular in the 1990s [1,2]. Denmark diet, only eating a controlled diet such as boiled eggs, black coffee, and grapefruit without salt or sugar, was a hit in the 2000s. In the 2010s, intermittent diets of fasting for a certain period
of time such as 12 hours a day or every other day was prevalent [3,4]. As such, dieting has been constantly changing over time. Even the target subject of diet has changed; switch from women-oriented diet to men-oriented one with the change of lifestyles. The market in Korea also expanded as the number of several types of diet increased, continuously growing from 200 billion won of total sales in the 2000s to 1.5 trillion won at 2010 and 3.2 trillion won at 2015 [5-7].

Attention is given to the diet of companion animals as well. The term ‘companion animals’, which brought a new perspective on animals that used to be treated like toys as family members or friends, was firstly brought by ethologist Konrad Lorenz in ‘The Human-Pet Relationship’ symposium held in Austria at 1983 and has been widely used since then [8,9]. In recent years, the demand for companion animals has been increasing due to several benefits from having them such as overcoming loneliness, preventing depression, and emotional communication of owners making up for the absence of family caused by single household, low fertility rate and increased aging population. Therefore, the companion animal market is also growing [10,11]. The market scale has increased from 900 million won at 2010 to 1.8 trillion won at 2015 as the food sector particularly taking a large portion. The marker expected to be grow up to approximately 6 trillion won by 2020 [12,13].

Selection and control of the feed for companion animal usually depend on the owner's living style and circumstances. Moreover, a walk or other outdoor activities of companion animals are usually performed with the owner. In other words, the owner’s attention to the companion animal influences the life pattern or undesirable health status of the animal such as obesity and nutritional imbalance. Obesity, in particular, is emerging as a serious problem. According to the result of Association for Pet Obesity Prevention in 2016 [14], 54% of dogs and 59% of cats are overweight or obese, indicating that the obesity level of companion animals must not be ignored. It has been discovered that obesity in companion animals can also lead to cardiovascular diseases and diabetes, causing various health problems [15]. Thus, a diet to prevent obesity is important for companion animals.

In this study diets of human, who can voluntarily follow a diet, and companion animals, which require care by human, were analyzed using big data algorithm to compare these diets. For the analysis, the sentences indicating diets of human and companion animals were collected from portal site Naver for 1 year from January 1, 2016 to December 31, 2016. Then the diets of human and companion animals were compared through diet related keyword frequency analysis, N-gram analysis (analysis on the frequency and direction of words that concurrently appeared with diet keywords), keyword network analysis (grouping similar words and finding the correlation between them) and seasonality analysis (analysis on appearance frequency of diet keywords by month).

**MATERIALS AND METHODS**

**Study subject and period**

The data for this study was collected from Naver, the best-known portal site in Korea, to analyze diets of human and companion animals. The sentences including the keyword ‘diet’ were collected for both human and companion animals. And only the data from January 1, 2016 to December 31, 2016 was collected.
Data collection and preprocessing

The Python 2.7 (Pycon, USA) program was used for collection of data, and the application programming interface (API) function in Naver search was used to gather sentences including the keyword ‘diet’ from Naver blogs, websites, news, and cafes (forums). Nouns related to a diet were extracted from the collected sentences using MeCab, which is the open source program by Eunjeon Hannip for morphological analysis designed for the Korean language with minimal alteration.

For data analysis, preprocessing on keywords that were classified into nouns was performed for data analysis using MeCab as follows. First, compound words formed with more than 2 words were converted back to single words. Second, words that have different forms but have same meaning such as ‘woman’ and ‘female’ were unified and converted to a single word.

Lastly, words expressed with postpositions and pronouns such as ‘by’, ‘with’ and ‘for’ were deleted as they could not be used for the analysis. Total 115,805 relevant keywords including ‘diet’ were found for human and the total frequency count of the words was 13,926,052. The total number of relevant keywords including ‘diet’ for companion animals was 61,015 and the total frequency count of the words was 2,113,544.

Analysis method

1) Keyword frequency analysis
   It is useful to analyze the frequency of the keywords associated with the subject keyword in descending order for intuitively understanding massive quantities of collected data [16] because the associated keywords considerably reflect the interests and attention about the subject keyword. Therefore, the keyword frequency analysis was used to find the keywords with high frequency of appearance from the data.

2) N-gram analysis
   This study conducted N-gram analysis because the keyword frequency analysis only analyzed the frequency of the entire data but not the co-appearance and concentration of the topic keyword and associated keywords. N-gram analysis evaluated the density of the words and their frequency by classifying n-number of syntactic words or syllables designated by the researcher [17]. Therefore, this study conducted N-gram analysis by separating keyword unit and the direction between keywords was discovered using the N-gram network word tree.

3) Keyword network analysis
   Keyword network analysis is a method of extracting keywords as a network and formats a connection between keywords by understanding and analyzing the relation between the keywords of noun and adjective forms in the sentence [18,19]. The network is composed of a link that shows the relations with the nodes represented by the actors [20]; a network is composed of links representing relationship and nodes represented by actors; keywords represent the nodes and the connection between keywords is shown as the link in this study. Therefore, overall network structure, forms and features of the link and nodes, direction and relations of the nodes, etc. can be visually expressed through keyword network analysis [21]. To enable data into social matrix for keyword network analysis, ‘Textom’ which is a social matrix program provided by The IMC Co., Ltd. was used. And this study used CONvergence of iterated CORrelations (CONCOR) analysis among keyword network analyses and the Ucinet 6 program to visualize the results.
CONCOR analysis form clusters after establishing relationships between similar keywords by understanding relation between the blocks and identifying the blocks of nodes according to the Pearson correlation of matrices [22].

4) Seasonality analysis  
It was considered that increased interests and attention in a diet lead to the increase in exposure frequency of the diet keyword, and therefore the diet keyword was analyzed by monthly appearance frequency for seasonality analysis.

RESULT

Keyword frequency analysis
The result of analyzing the appearance frequency of keywords associated with diets for human and companion animals is shown in Table 1. The keywords associated with diets for human were shown in order of ‘exercise’ (n = 191,032), ‘eating’ (n = 102,631) and ‘health’ (n = 86,199). The keywords associated with diets for companion animals were shown in order of ‘dog’ (n = 41,629), ‘feed’ (n = 31,068) and ‘cat’ (n = 30,476).

N-gram analysis
N-gram analysis for analyzing the co-appearance frequency through the density between keywords was conducted on human and companion animals. As the result shown in Table 2, the keyword that appeared the most concurrently with ‘diet’ was ‘menu’ with 59,376 counts followed by ‘diet’ and ‘start’ keywords with 41,594 counts. ‘Diet’ and ‘success’ keywords appeared 37,964 times. On the other hand, ‘diet’ and ‘method’ keywords appeared the most with 6,660 counts for companion animals which were followed by ‘diet’ and ‘feed’ keywords with 5,689 counts. ‘Diet’ and ‘effect’ keywords appeared 3,345 times.

Table 1. The frequency of diet keywords for human and companion animal related to the diet by simple frequency analysis in 2016

<table>
<thead>
<tr>
<th>Rank</th>
<th>Human</th>
<th>Frequency</th>
<th>Companion Animal</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Exercise</td>
<td>191,032</td>
<td>Dog</td>
<td>59,376</td>
</tr>
<tr>
<td>2</td>
<td>Menu</td>
<td>102,631</td>
<td>Feed</td>
<td>41,594</td>
</tr>
<tr>
<td>3</td>
<td>Health</td>
<td>86,199</td>
<td>Cat</td>
<td>37,964</td>
</tr>
<tr>
<td>4</td>
<td>Tea</td>
<td>85,602</td>
<td>Companion Animal</td>
<td>33,363</td>
</tr>
<tr>
<td>5</td>
<td>Start</td>
<td>80,826</td>
<td>Method</td>
<td>27,602</td>
</tr>
<tr>
<td>6</td>
<td>Effect</td>
<td>79,405</td>
<td>Obesity</td>
<td>24,082</td>
</tr>
<tr>
<td>7</td>
<td>Breakfast</td>
<td>65,966</td>
<td>Snack</td>
<td>21,261</td>
</tr>
<tr>
<td>8</td>
<td>Today</td>
<td>62,117</td>
<td>Health</td>
<td>22,996</td>
</tr>
<tr>
<td>9</td>
<td>Moment</td>
<td>60,327</td>
<td>Exercise</td>
<td>22,956</td>
</tr>
<tr>
<td>10</td>
<td>Success</td>
<td>57,181</td>
<td>Recommend</td>
<td>22,345</td>
</tr>
<tr>
<td>11</td>
<td>Diary</td>
<td>53,780</td>
<td>Effect</td>
<td>21,485</td>
</tr>
<tr>
<td>12</td>
<td>Management</td>
<td>50,154</td>
<td>House</td>
<td>18,849</td>
</tr>
<tr>
<td>13</td>
<td>Method</td>
<td>49,854</td>
<td>Human</td>
<td>17,329</td>
</tr>
<tr>
<td>14</td>
<td>Kg</td>
<td>48,450</td>
<td>Kg</td>
<td>15,956</td>
</tr>
<tr>
<td>15</td>
<td>Farewell</td>
<td>47,735</td>
<td>Today</td>
<td>15,279</td>
</tr>
<tr>
<td>16</td>
<td>Review</td>
<td>45,772</td>
<td>Belly</td>
<td>15,207</td>
</tr>
<tr>
<td>17</td>
<td>Dinner</td>
<td>45,452</td>
<td>Review</td>
<td>15,090</td>
</tr>
<tr>
<td>18</td>
<td>Body</td>
<td>45,282</td>
<td>Food</td>
<td>13,360</td>
</tr>
<tr>
<td>19</td>
<td>Thinking</td>
<td>43,200</td>
<td>Start</td>
<td>12,338</td>
</tr>
<tr>
<td>20</td>
<td>Reduction</td>
<td>42,438</td>
<td>Success</td>
<td>11,890</td>
</tr>
</tbody>
</table>
The direction between keywords was found through Keyword1 and Keyword2. As a result, the keyword 'diet' firstly appeared by Keyword1 and the keyword 'eating' was appeared by Keyword2. In other words, the human being search the menu first when they are going on a diet. In case of companion animals, the keyword 'method' appears after the keyword 'diet' is searched first. And the keyword 'diet' appears again and then 'feed'. Both human and companion animals showed less frequency of 'diet' from Keyword1 and 'exercise' from Keyword2 compared to N-gram. Therefore, human and companion animals commonly have a higher interest in dietary control (human: menu, companion animals: feed) for a diet.

**Keyword network analysis**

This study used CONCOR analysis for keyword network analysis. Figure 1 shows keyword network analysis of diets for human which are classified into total four groups. Keywords including 'menu', 'breakfast', 'lunch', 'dinner', 'snack', and 'meal' are linked into a food group. Keywords including 'exercise', 'fitness', 'body shape', 'yoga', 'health', 'summer', 'boxing', and 'recommendation' are linked into an exercise group. Keywords including 'salad', 'chicken breast', 'taste', 'tea', and 'price' are linked into a commercial diet food group. Keywords including 'JUVIS', 'review', 'clinic', 'success', and 'care' are linked into a commercial diet program group. Figure 2 shows keyword network analysis of diets for companion animals which is classified into total three groups. Keywords including 'feed', 'chicken breast', 'snack', 'food', and 'meal' are linked into a food group. Keywords including 'walk', 'belly', 'exercise', 'care', and 'home' are linked into an exercise group. Keywords including 'hospital', 'medical care', 'surgery', 'treatment', 'price', and 'success' are linked into a professional medical support group.

**Seasonality analysis**

To find out the season during which diet keywords were searched the most, the frequency of diet keywords was divided into monthly period and analyzed. According to Figure 3, human and companion animals both had the similar monthly frequency of diet keywords. Human showed the highest frequency in October with 144,584 counts which was followed by 143,842 counts.
The Analysis of a Diet for the Human and the Companion Animal using Big Data

Figure 1. Keyword network analysis related to the human diet in 2016

Figure 2. Keyword network analysis related to the companion animal diet in 2016
counts in July. Similar to human, companion animals showed the highest frequency in July with 26,163 counts which was followed by 24,064 counts in September. Human and companion animals both showed increasing frequency of diet keywords from February to July.

**DISCUSSION**

A diet is a main interest in human being and diets for companion animals also get much attention recently. Therefore, this study conducted keyword frequency analysis, N-gram analysis, keyword network analysis, and seasonality analysis through big data after collecting sentences that represent diets for human and diets for companion animals from the portal site Naver.

In keyword frequency analysis, human showed 191,032 counts of the keyword 'exercise'. This high frequency was due to the repeated keyword 'exercise' found in sentences such as ‘Exercises helpful for a diet include aerobic exercise, weight training exercise…’. In the study by Jung and Chang [23], the keyword ‘menu’ was counted as 79,675 at 2015 but it marked 102,631 at 2016, which has increased by 22,956 compared to the previous year indicating that human's interests in dietary menu have increased.

On the other hand, companion animals showed 41,629 counts of the keyword 'dog' and 30,476 counts of the keyword 'cat'. Hwang and Kim [12] published that 4.4 million dogs and 1.16 million cats are raised at home as of 2012. There are more dogs than cats raised at home, and therefore attention on a diet is higher for dogs than cats. Unlike human, companion animals showed keywords like ‘human’ and ‘home’, and this is because diets of companion animals are carried out by the owners, human beings and they are looking for diet methods that can be tried at home.

To complement the limit of keyword frequency analysis that can only enumerate the simple frequency of appearance from the entire data, the result of diets for human and companion animals was analyzed through N-gram analysis which can find the density between keywords. Diet-menu (59,376 counts) appeared the most for human which was followed by diet-tea (23,261 counts), diet-food (22,956 counts), and diet-lunchbox (12,138 counts). For
companion animals, results for eating were shown in order of diet-pet food (5,989 counts), diet-menu (2,472 counts), dog-pet food (2,194), cat-pet food (2,152), homemade-snack (2,118 counts), diet-food (1,592 counts), and pet food-recommendation (1,521 counts). Bland et al. [24] published a study on eating and obesity, and found that obesity of dogs was associated with eating. Jones and Lewis [25] judged that feed was important for the diet of dogs and developed pet food was effective for weight loss. Accordingly, eating has a significant impact on a diet of companion animals as well as human.

Exercise (33,363 counts) was shown less than the N-gram analysis frequency related to eating. Mozaffarian et al. [26] published that weight loss through reduced eating is more effective than weight loss by exercise. According to such result, it is possible to expect that human being will have more interest in eating than exercise when attempting a diet and therefore the number of N-gram frequency of eating was higher than the frequency exercise. Similarly, companion animals showed lower frequency of diet-exercise (1,482 counts) than the N-gram frequency related to eating. Carciofi et al. [27] published that weight loss of companion animals depend on the cooperation of human being. Weight loss methods effective in human being might be applied to companion animals and therefore eating was shown higher than exercise.

This study aimed to find out which keyword could be grouped and associated with other keyword through the keyword network analysis. Keyword network analysis for diets of human was composed of total 4 groups; food group, exercise group, commercial diet food group, and commercial diet program group. This classification is identical to the diet keyword network analysis by Jung and Chang [23] in 2015 and 2016 whose study results also demonstrated 4 groups. But keywords under the groups have changed slightly. The keyword ‘tea’ was excluded from the food group of 2015 and the keyword ‘water’ was added to the 2016 group. The keywords ‘shake’ and ‘lunchbox’ were excluded from the commercial diet food group of 2015, and the keyword ‘salad’ was added to the 2016 group.

Groups classified for diets of companion animals were similar to those of human but there were some differences. Diets of companion animals did not create a commercial diet group but created a professional medical support group. Bland et al. [24] and Kang [15] reported factors that can cause obesity of companion animals include eating, heredity, breed, and neutralization status. Therefore, it seems that the profession medical support group was formed because the owner needs professional medical help if the obesity is caused by other factors than eating.

Keywords including ‘breakfast’, ‘lunch’, ‘dinner’, and ‘snack’ were found in the food group of human but no identical keywords except for ‘snack’ were found in the food group of companion animals. This finding is related to the fact that companion animals only take care of eating through feeds and snacks unlike human who take various forms of meals. Therefore, the owner must control many variables for dietary regulation, but the control variable of eating for companion animals is restricted to feeds and snacks.

Lastly, seasonality analysis was conducted through the monthly appearance frequency of diet keywords for human and companion animals, and both showed similar characteristics in graphs. Both subjects show an increasing trend from February to July, and the graphs once again rose in September with Chuseok (Korean Thanksgiving). Because clothing changes from heavy to light and reveals the body figure when the weather changes from
winter to spring and summer, which inspires eating [28,29]. Food from Chuseok stimulates an appetite for food and people tend to lose self-control and consume a lot of food during Chuseok. Chuseok food mostly consists of high-calorie meals [30,31]. It is considered that the appearance frequency of a diet increases in September with Chuseok. People’s motivation for a diet has an influence on the diet of pets and therefore the diet graphs of people and pets have similar aspects.

However, unlike companion animals, the diet graph of human shows the highest records in January and October. Diet fever hit in January as human start a new year. This finding seems to be related to the inspiration by successful diet stories of celebrities introduced through mass media including ‘Open List’ by tvN that covered the story of Yumin Noh who lost 30 kg and ‘Human Docu: Good People’ by MBC that covered the story of Seungyeon Sohn who lost 10 kg. In October, a “low-carb, high-fat” diet for restricting the consumption of carbohydrate and taking fat as the source of energy to lose weight attracted a people’s attention [32]. The diets of human are affected by seasons as well as mass media and trends, but the diets of companion animals are unaffected by these factors.

This study analyzed diets of human being and diets of companion animals by big data analysis, and confirmed that both human being and companion animals have high interests in eating when they are going on a diet. Since companion animals cannot plan their diets for themselves and diets are carried out by the owner, an effective diet plan for the owner is typically applied to animals and therefore the starting point of their diet is similar to the diet of the owner. Ultimately, the diet of companion animals depends on their owners. Acknowledging a diet as an interactive activity between human beings and companion animals rather than separating it for human beings and companion animals can enhance the benefit of the diet for companion animals.

**CONCLUSION**

This study used big data analysis to understand trends on diets of human and companion animals. Sentences including diet keywords were collected through the portal site Naver from January 1, 2016 to December 31, 2016. Keyword frequency analysis, N-gram analysis, keyword network analysis, and seasonality analysis were conducted on these sentences and the results are as follows.

1. The keyword that appeared the most in human through keyword frequency analysis was ‘exercise’ with 191,032 counts, and it was ‘dog’ in companion animals with 41,629 counts. The keyword ‘menu’ was shown high in human with 79,675 counts and the keyword ‘feed’ was shown high in companion animals with 31,068 counts.
2. In N-gram analysis that analyzed the link frequency between keywords, keywords related to eating were shown in order of diet-menu (59,376 counts), diet-tea (23,261 counts), diet-food (22,956 counts), and diet-lunchbox (12,138 counts) for human. For companion animals, keywords were shown in order of diet-feed (5,989 counts), diet-menu (2,472 counts), dog-feed (2,194 counts), cat-feed (2,152 counts), homemade-snack (2,118 counts), and diet-food (1,592 counts). The result related to diet-exercise in human (33,363 counts) and diet-exercise in companion animals (1,482 counts) was shown low frequency compared to eating. Thus, it was confirmed that human beings and companion animals both prefer diet over exercise.
3. As a result of conducting keyword network analysis that grouped keywords associated with a diet with similar characteristics, humans were classified into a total of 4 groups and companion animals were classified into a total of 3 groups. The food group and the exercise group were commonly shown in both human and companion animals, but human had the commercial diet program group and the commercial diet food group whereas companion animals had the professional medical support group.

4. The seasonality analysis showed that human and companion animals shared similar graphs. The starting time of diets in human and companion animals were similar, but human’s diets were significantly affected by mass media and trend as in January and October.

In conclusion, the diets of human and companion animals showed similar tendencies, particularly having higher preference for dietary control. The diets of companion animals are determined by the choice of their owners because, typically, diet methods effective in owners are applied to companion animals. However, in Korea, there is no study of the correlation between obesity of human and companion animals; therefore, it is necessary to have empirical demonstration of whether a correlation of obesity between human and companion animals exists.

REFERENCES


https://doi.org/10.7762/cnr.2017.6.4.256