Introduction

In 2021, the amount of online information related to indoor crop cultivation increased by about 57% compared to the previous year, and in particular, the amount of mentions of balcony vegetable gardens, one of the indoor gardening trends, increased by 151% compared to 2019 (EPIS, 2021). As a result of the Korea Housing Survey in 2020, the percentage of households living in apartments among general households in Korea was 51.1% (MOLIT, 2021), suggesting the need to pay attention to home gardening, especially balcony vegetable gardens, since more than half of the housing types are apartments. Moreover, the mental health problems caused by the unprecedented COVID-19 pandemic have become a global disaster, warning that mental health services and support must be strengthened (Brunier and Drysdale, 2022). In this situation, studies verifying the applicability of home gardening as a solution are constantly conducted.
Sia et al. (2022) stated that home gardening is an effective method to promote interaction with nature and mental resilience in people living in densely populated cities. Theodorou et al. (2021) claimed that gardening activities reduce stress and help promote mental health, and Howarth et al. (2020) reported that gardening activities improve mental well-being, such as increased physical activity and decreased social isolation. Egerer et al. (2022) confirmed that through home gardening, it is possible to secure food safety and gain not only economic benefits through self-supply of food, but also mental health and benefits such as connecting with nature and relieving stress.

As such, home gardening contributes to realizing not only emotional stability, but also physical health and self-sufficiency (Gerdes et al., 2020; Nicola et al., 2020) through self-grown healthy food in times of all-pervading anxiety and threat. Moreover, since the outbreak of COVID-19, the mention of feeling the value of healing in daily life through balcony vegetable gardens has increased (EPIS, 2021), which supports the argument in various studies that mental health problems including psychological stress of high-rise apartment residents can be reduced by creating green space (WHO, 2015; WHO, 2017; Larcombe et al., 2019).

With the public health crisis and the growing interest in the residential environment, the climate crisis and the ongoing COVID-19 emphasize the importance of green (Pamukcu-Albers et al., 2021). Various studies have been conducted, such as a study on gardening of home-style vegetable gardens in Canada in response to the COVID-19 pandemic reflecting this social climate (Mullins et al., 2021), a study on the role of pocket gardens at home and their correlation with quality of life during the pandemic (Sanusi et al., 2021) and a study that identified the relationship between urban gardening and well-being in the COVID-19 pandemic using social and environmental factor approaches (Harding et al., 2022).

However, there is insufficient research on home gardening from the outbreak of COVID-19 in Korea. In particular, there is no research identifying public awareness and trend about balcony vegetable gardens as a home gardening activity in apartments that take up half of the housing types in Korea, despite the increasing public interest and mentions about balcony vegetable gardens. Therefore, this study was conducted to provide basic data to activate related research and set the policy direction by analyzing the trend in balcony vegetable gardens in the last 4 years including before and after COVID-19 using big data.

**Research Methods**

**Data extraction and cleaning**

Big data analytics discovers hidden meanings and patterns in the context of data, and interest in extractable values and utilization methods is increasing worldwide (Jung, 2010; Lee and Jeon, 2016). In particular, big data analytics is suitable for examining various perceptions and phenomena that occur in society as a whole rather than in a specific group (Choi and Choi, 2016). This study used big data collection and analysis solution TEXTOM v5.0 (The IMC, Korea) as a data extraction tool to analyze the trend through big data produced with active participation of general users such as blogs or cafes (Yoon et al., 2009).

First, data for 4 years from January 1, 2018 to December 31, 2021, posted with ‘balcony vegetable garden’ as a keyword were collected from blogs and cafes of Naver, Korea’s leading portal site with the largest number of users. Among them, the dates of 5,011 cases of valid data were confirmed, and 2,000 cases from January 1, 2018 to December 21, 2019 (before the outbreak of COVID-19) and 3,011 cases from January 1, 2020 to December 31, 2021 (after the outbreak of COVID-19) were used for final analysis. For morpheme refinement, Okt (Open Korean Text) morphological analysis of the KoNLPy package widely used for Korean language information processing in natural language processing (NLP) of Python 3.9.5 (Pycon, USA) was used, tokenizing nouns with actual meanings in sentences as free morphemes.

Okt morphological analyzer is an open-source-based Twitter morphological analyzer that uses the normalize function to normalize words that are not listed in the Standard Korean Language Dictionary, and it is known to be quicker than other analyzers and extract words relatively accurately and consistently (Lee et al., 2021; Hur et al., 2022; Sun and Lim, 2019). Therefore, this study adopted the Okt ana-
lyzer and converted abbreviations such as ‘C.T. (cherry tomato)’ and new words or compounds such as ‘pet plant’ into one word and added to the user dictionary, while also conducting the preprocessing process to designate and eliminate stop words such as postpositions, special characters, and conjunctions.

**Analysis method (data analysis)**

To identify keywords related to balcony vegetable gardens, the term frequency-inverse document frequency (TF-IDF) weights were calculated using TfidfVectorizer of the Scikit-learn library and the data were converted (Sun and Lim, 2019). TF-IDF is a method of quantifying how important a specific word is in a document (Woo and Kim, 2020). Through the extracted keywords, the overall trend and characteristics of the document can be identified.

Next, topic modeling, which is a probabilistic model algorithm (Blei et al., 2003) for extracting topics from a corpus of unstructured text data, was used as a text mining technique. Latent Dirichlet Allocation (LDA) is one of the topic modeling techniques that are widely used, and it is known to have clear topic extraction methods by modeling a document as a topic, which is a collection of words that make up a document, and a mixture of topics (Jeong and Park, 2021), and it is useful for online text analysis since it is easy to find the relation between meanings (Nam, 2016).

This study used the LDA topic modeling algorithm of the Gensim package to conduct LDA topic modeling by setting the hyper parameters as $\alpha = 0.1$, $\beta = 0.01$, and iterations = 1,000 (Rashid et al., 2019; Lee et al., 2020, Kim, 2020). $\alpha$ controls how dense the distribution of topics in the literature will be, and $\beta$ controls how dense the distribution of words in the topic will be (Han and Heo, 2021). Next, by measuring the coherence index of the topics, the adequate number of topics K was determined. Higher coherence index indicates that each topic of the LDA analysis is consistent and is composed of similar words and well reflects the actual literature (Han and Heo, 2021; Cho, 2022). The analysis above was conducted through Python (Fig. 1).

**Results and Discussion**

**Keyword analysis (TF-IDF analysis)**

This study divided 5,011 cases of data for 4 years from January 1, 2018, extracted with balcony vegetable garden as a keyword, into 2 years before the outbreak of COVID-19 (Jan. 1, 2018-Dec. 31, 2019) and 2 years after the outbreak (Jan. 1, 2020-Dec. 31, 2021), calculated TF-IDF values, and came up with the main keywords. TF-IDF gives weights to simple frequency to represent how much weight it has within a document, thereby enabling users to identify the importance more accurately (Kwak and Kim, 2019).

Table 1 shows the keyword analysis results of the top 30 TF-IDF values of the two groups before and after the outbreak of COVID-19, excluding the search keyword ‘balcony vegetable garden’. First, before the outbreak of COVID-19, the keyword showing the highest importance was seedling (1432.33), followed by lettuce (1302.14), harvest (1224.23), vegetable (838.93), and seed (763.89). During COVID-19, the keyword showing the highest importance was harvest (2090.19), followed by vegetable (1825.06), healing (1797.66), my place (1434.12), and hobby (1318.74). The words ranked at the top can be seen as keywords representing the public awareness or behavioral characteristics toward balcony vegetable gardens.

For the top 30 keywords before the outbreak of COVID-19,
words such as flowerpot (744.26), child (633.08), fruit (621.64), kidney bean (612.44), experience (610.39), school (555.74), Arbor Day (555.42), and day care center (525.15) showed that balcony vegetable gardens are created through institutions carrying out home gardening programs and experience activities.

Keywords derived newly after the outbreak of COVID-19 compared to before were urban farmer (1235.64), mother (1175.50), kit (1116.82), pet plant (1097.83), stay at home (947.97), self-sufficiency (823.06), COVID-19 (813.51), and Daiso (756.16). These words implied that home gardening began to receive attention as a healthy hobby with people spending more time at home due to COVID-19 and that people are healing themselves by participating in balcony vegetable gardening with kits that are easy to purchase at stores selling daily necessities like Daiso.

**Topic modeling**

This study conducted topic modeling (Kim and Kim, 2020), which is effective in summarizing a large amount of text data by forming a cluster of words in a document set to understand the main topics and inherent meanings of posts related to balcony vegetable gardens. First, the coherence of topics was measured, with higher coherence scores indicating that each topic has coherence (Kim and Kim, 2019). In this study, the coherence index was highest at 0.356 when there were 4 topics before the COVID-19 outbreak, and it was highest at 0.493 when there were 5 topics during COVID-19, based on which this study conducted the analysis by selecting the optimal number of topics (Fig. 2). Next, 10 keywords were derived for each of 4 topics before COVID-19 and 5 topics after COVID-19, giving name to each topic (Table 2).

The results of naming the four topics before the outbreak of COVID-19 are as follows. Top keywords of Topic 1 included 'vegetable', 'lettuce', 'cherry tomato', 'crop', 'perilla leaf', 'sprout', 'cultivation', 'kind', 'urban agriculture', and 'taste'. These words helped identify the trend of cultivating various kinds of leafy vegetables through balcony vegetable gardens as part of urban agriculture. Accordingly, Topic 1 was named 'main crops of balcony vegetable gardens'.

Top keywords of Topic 2 included 'flowerpot', 'kidney bean', 'experience', 'soil', 'weekend', 'seedling', 'education', 'child', 'day care center', and 'school', which showed the institutions in which horticultural experience took place as well as the contents of such activities. They also confirmed that horticultural activities continued from educational institutions to home. Accordingly, Topic 2 was named 'gardening experience and inflow path of home gardening'.

Top keywords of Topic 3 included 'seed', 'herb', 'weed',

---

**Table 1. Top 30 Keywords before and After the COVID-19 pandemic**

<table>
<thead>
<tr>
<th>Rank</th>
<th>Keywords</th>
<th>TF-IDF</th>
<th>Before COVID-19</th>
<th>Rank</th>
<th>Keywords</th>
<th>TF-IDF</th>
<th>After COVID-19</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Seedling</td>
<td>1432.33</td>
<td>Perilla leaf</td>
<td>16</td>
<td>Green</td>
<td>1037.49</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Lettuce</td>
<td>1302.14</td>
<td>School</td>
<td>17</td>
<td>Daily life</td>
<td>1020.17</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Harvest</td>
<td>1224.23</td>
<td>Arbor Day</td>
<td>18</td>
<td>Plant butler</td>
<td>1006.61</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Vegetable</td>
<td>838.93</td>
<td>Day care center</td>
<td>19</td>
<td>My place</td>
<td>997.30</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Seed</td>
<td>763.89</td>
<td>Crop</td>
<td>20</td>
<td>Fun</td>
<td>982.40</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Flowerpot</td>
<td>744.26</td>
<td>Chives</td>
<td>21</td>
<td>Basil</td>
<td>968.12</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Cherry tomato</td>
<td>697.04</td>
<td>Compact</td>
<td>22</td>
<td>Interest</td>
<td>965.97</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Child</td>
<td>633.08</td>
<td>This year</td>
<td>23</td>
<td>Stay at home</td>
<td>947.97</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Fruit</td>
<td>621.64</td>
<td>Education</td>
<td>24</td>
<td>Vitality</td>
<td>938.35</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Kidney bean</td>
<td>612.44</td>
<td>Fertilizer</td>
<td>25</td>
<td>Roof top</td>
<td>829.17</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Experience</td>
<td>610.39</td>
<td>Bug</td>
<td>26</td>
<td>Self sufficiency</td>
<td>823.06</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Apartment</td>
<td>596.36</td>
<td>Weekend</td>
<td>27</td>
<td>Kit</td>
<td>813.51</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Flower</td>
<td>592.42</td>
<td>Herb</td>
<td>28</td>
<td>COVID-19</td>
<td>809.17</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Raised garden bed</td>
<td>591.54</td>
<td>Sunlight</td>
<td>29</td>
<td>Joy</td>
<td>777.89</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Root</td>
<td>585.30</td>
<td>Soil</td>
<td>30</td>
<td>Daiso</td>
<td>756.15</td>
<td></td>
</tr>
</tbody>
</table>

450 • Journal of People, Plants, and Environment Vol. 25, No. 5, 2022
'sharing', 'preparation', 'flower', 'germination', 'fertilizer', 'bug', and 'repotting'. These keywords represent horticultural activity itself, proving that users were actively participating in home gardening activities such as seeding on the apartment balcony, weeding or removing bugs, and repotting. Moreover, these words showed the trend of sharing a set of procedures for balcony vegetable gardening activities. Accordingly, Topic 3 was named 'balcony vegetable gardening activities'.

Top keywords of Topic 4 included 'raised garden bed', 'apartment', 'set', 'rooftop', 'harvest', 'compact', 'salad', 'shelf', 'space', and 'recycling', which showed the types of vegetable gardens preferred by those participating in balcony vegetable gardening activities. They were mostly creating balcony vegetable gardens in compact space using raised garden bed sets or flowerpots. In addition, balcony vegetable gardens using various recyclables such as 'PET bottle' and 'Styrofoam' were also receiving attention. Accordingly, Topic 4 was named 'types of balcony vegetable gardens'.

The results of naming the five topics during the outbreak of COVID-19 are as follows. Top keywords of Topic 1 included 'my place', 'COVID-19', 'daily life', 'stay at home', 'healing', 'first time', 'hydroponics', 'nature', 'indoor', and 'vegetable'. Due to continued social constraints caused by COVID-19, balcony vegetable gardening at home became a part of daily life, through which people were healing themselves. Accordingly, Topic 1 was named 'daily life and home gardening during COVID-19'.

**Table 2. Results of LDA topic modeling**

<table>
<thead>
<tr>
<th>Period</th>
<th>Keywords</th>
<th>Title of topics</th>
<th>Proportion (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Before COVID-19 (2018.01.01. - 2019.12.31.)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Topic 1</td>
<td>Vegetable, Lettuce, Cherry tomato, Crop, Perilla leaf, Sprout, Cultivation, Kind, Urban agriculture, Taste</td>
<td>Main crops of Balcony vegetable garden</td>
<td>28.4</td>
</tr>
<tr>
<td>Topic 2</td>
<td>Flowerpot, Kidney bean, Experience, Soil, Weekend, Seedling, Education Child, Day care center, School</td>
<td>Gardening experience and inflow path of home gardening</td>
<td>31.0</td>
</tr>
<tr>
<td>Topic 3</td>
<td>Seed, Herb, Weed, Sharing, Preparation Flower, Germination, Fertilizer, Bug, Repotting</td>
<td>Balcony vegetable garden activities</td>
<td>17.9</td>
</tr>
<tr>
<td>Topic 4</td>
<td>Raised garden bed, Apartment, Set, Rooftop, Harvest, Compact, Salad, Shelf, Space, Recycling</td>
<td>Types of balcony vegetable garden</td>
<td>22.7</td>
</tr>
<tr>
<td><strong>After COVID-19 (2020.01.01. - 2021.12.31.)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Topic 2</td>
<td>Green, Raised garden bed, Organic, Sustainability, Seedling, Apartment, Climate action, Distribution, Sprout, Seed, Environment</td>
<td>Balcony vegetable garden and environment practices</td>
<td>14.6</td>
</tr>
<tr>
<td>Topic 3</td>
<td>Fun, Weekend, Plant butler, Gratitude, Vitality, Joy, Mind, Growth, Expectation, Success</td>
<td>Balcony vegetable garden and emotions</td>
<td>17.3</td>
</tr>
<tr>
<td>Topic 4</td>
<td>Urban farmer, Mother, Basil, Daiso, Green onion, Challenge, Posting, Kit, Interest, Food</td>
<td>Inflow path of Balcony vegetable garden activities</td>
<td>23.8</td>
</tr>
<tr>
<td>Topic 5</td>
<td>Self sufficiency, Beginner, Health, Creation, Pet plant, Harvest, Flowerpot, Picture, Hobby, Recommendation</td>
<td>Balcony vegetable garden and hobbies</td>
<td>17.8</td>
</tr>
</tbody>
</table>
Top keywords of Topic 2 included ‘green’, ‘raised garden bed’, ‘organic’, ‘sustainability’, ‘seedling’, ‘apartment’, ‘climate action’, ‘distribution’, ‘sprout’, ‘seed’, ‘and environment’. These words show the trend of sharing organic food people grew themselves, as well as the trend of aiming for environmental practices and sustainability such as climate action. Accordingly, Topic 2 was named ‘balcony vegetable gardens and environmental practices’.

Top keywords of Topic 3 included ‘fun’, ‘weekend’, ‘plant butler’, ‘gratitude’, ‘vitality’, ‘joy’, ‘mind’, ‘growth’, ‘expectation’, and ‘success’. They expressed the emotions and feelings people get from balcony vegetable gardening. In particular, there were many words expressing positive emotions, which showed the emotional support that people obtain from home gardening activities. Accordingly, Topic 3 was named ‘balcony vegetable gardens and emotions’.

Top keywords of Topic 4 included ‘urban farmer’, ‘mother’, ‘basil’, ‘Daiso’, ‘green onion’, ‘challenge’, ‘posting’, ‘kit’, ‘interest’, and ‘food’. These words showed interest in food and how people began to find interest in balcony vegetable gardens through family or posts and actually attempting to participate in horticultural activities by raising seeds or seedlings purchased at a store selling daily necessities. Accordingly, Topic 4 was named ‘inflow path of balcony vegetable gardening activities’.

Top keywords of Topic 5 included ‘self-sufficiency’, ‘beginner’, ‘health’, ‘creation’, ‘pet plant’, ‘harvest’, ‘flowerpot’, ‘picture’, ‘hobby’, and ‘recommendation’. This showed the trend that balcony vegetable gardening is receiving attention as a hobby for a healthy life, and the trend of new words like pet plant as well as interest in the activity. Accordingly, Topic 5 was named ‘balcony vegetable gardens and hobbies’.

The results above imply that balcony vegetable gardening activities were encouraged and received attention as a type of urban agriculture before COVID-19, while after COVID-19 they are established as activities for healing, emotional support, environmental campaign, and hobby, and that this therapeutic effect and expected outcome may be a significant key to the promotion and development of urban agriculture. In particular, taking care of plants through home gardening like balcony vegetable gardens affects emotional changes such as sense of achievement, self-esteem, and security, thereby promoting individual welfare and environmental benefits (Park and Shin, 2021; Harding et al., 2022).

Among urban agriculture, balcony vegetable gardening activities have the advantage of enabling people to cultivate crops at home without moving, check the growth, have fun, and tend the vegetable garden regardless of time and place (Kim and Yoon, 2016). The revitalization of urban agriculture based on these advantages is considered to be a solution to the social problems derived from the public health crisis caused by the infectious disease called COVID-19. Furthermore, it can be inferred that balcony vegetable gardens will become more popular as there is a growing interest in safe food supply and environmental practices through self-sufficiency.

Conclusion

It has been reported that the use of balcony vegetable gardens is increasing along with the constant increase in the population living in apartments among housing types in Korea, and these home gardening activities are expected to improve mental, physical, and social health (Soga et al., 2017; Lal, 2020; Lin et al., 2021). In accordance with the increasing mentions and usage rates of balcony vegetable gardens, this study was conducted to provide basic data for collecting big data using balcony vegetable gardens as the keyword, analyzing the trends, and effectively applying them to related policies and research.

First, 5,011 cases of data extracted were divided into 2 years before the outbreak of COVID-19 (Jan. 1, 2018-Dec. 31, 2019) and 2 years after (Jan. 1, 2020-Dec. 31, 2021), and main keywords were derived by calculating TF-IDF values. Next, LDA topic modeling was conducted, and the four topics before the outbreak of COVID-19 were named ‘main crops of balcony vegetable gardens’, ‘gardening experience and inflow path of home gardening’, ‘balcony vegetable gardening activities’, and ‘types of balcony vegetable gardens’. The five topics after the outbreak of COVID-19 were named ‘daily life and home gardening during COVID-19’, ‘balcony vegetable gardens and sustainability’, ‘balcony vegetable gardens and emotions’, ‘balcony vegetable gar-
dens and hobbies', and 'inflow path of balcony vegetable gardening activities'.

Based on the topic modeling results obtained in this study, the trends in balcony vegetable gardens before and after the outbreak of COVID-19 can be summarized as follows. First, vegetable gardening activities, which began to receive attention as part of urban agriculture such as growing leafy vegetables on the balcony of an apartment, gradually began to expand into a healthy indoor hobby after the outbreak of COVID-19. In particular, there was a trend in which people form positive emotions and heal through balcony vegetable gardening activities as a healthy hobby that improves physical and mental health and actively share their experiences and daily life online.

Second, it was possible to identify the change in the inflow path of participation in balcony vegetable gardening activities. The inflow of participants before COVID-19 had been through educational institutions such as schools or daycare centers as well as outdoor horticultural activities that led to home. On the other hand, after COVID-19, as people showed more interest in health and self-sufficiency, they obtained information from families or online posts and participated voluntarily. This shows the importance of using online platforms in providing information about related activities or programs to revitalize and develop urban agriculture.

Third, unlike before when the focus had been only on balcony vegetable gardening itself, the trend after COVID-19 showed that people aimed for a sustainable life through balcony vegetable gardening. This implies that balcony vegetable gardening is perceived as an activity for environmental practices such as climate action, crop distribution, and sustainability, and this reflects the values and expected outcomes the public has toward urban agriculture. This result will help establish policies and set the direction for urban agriculture in the future.

This study has significance in that it identified the trends in balcony vegetable gardens before and after the outbreak of COVID-19 based on big data and analyzed the inherent meanings. However, it has the following limitations. First, the results cannot be generalized since the data collection platform was limited to just Naver alone, thereby not using data from various platforms. By collecting data from all kinds of social media known to be used by various age groups or Google that shows a constant growth in users, it would be possible to analyze the trends with greater objectivity based on more massive data.

Second, the data collection period was limited to the last 4 years from 2018 to 2021. By adding data from 2013 to 2017 that is the first 5-year Urban Agricultural Fostering Plan and comparing the first and second plans, more detailed changes in the trends can be observed.

Third, this study used only the Okt morphological analyzer of KoNLPy, which is the Korean language analysis package of Python. Despite the benefits of the Okt analyzer such as normalization function and quick processing, it does not operate properly without directly assigning new words or proper nouns not registered on the library dictionary. Since it is possible to use various Korean language morphological analyzers together with the development of natural language technology (Kim et al., 2021), future studies must improve the accuracy of morphological analysis with integrated use of multiple analyzers such as MeCab, Komoran, Kkma, and Hannanum of KoNLPy.

References


Woo, Y.H. and H.H. Kim. 2020. Topic analysis of the national petition site and prediction of answerable petitions based on deep learning. KIPS Transactions on Software and
Trend Analysis of Balcony Vegetable Gardens in Korea, Before and After COVID-19 Pandemic Using Big Data

Data Engineering 9(2):45-52. https://doi.org/10.3745/KTSDE.2020.9.2.45