Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Travel patterns of free-floating e-bike-sharing users before and during COVID-19 pandemic

Seung Eun Choi\textsuperscript{a}, Jinhee Kim\textsuperscript{b,}\textsuperscript{*}, Dayoung Seo\textsuperscript{c}

\textsuperscript{a} School of Civil and Environmental Engineering, Transportation Systems Engineering, Georgia Institute of Technology, 790 Atlantic Drive, Atlanta, GA 30332, United States of America
\textsuperscript{b} Department of Urban Planning and Engineering, Transportation Planning, Yonsei University, Seoul 03722, South Korea
\textsuperscript{c} Product Team Lead, Elecle, Seoul 06611, South Korea

\textbf{ARTICLE INFO}

\textbf{Keywords:}
Free-floating micro-mobility
Dock-less
Electric-bike sharing
Pandemic
Clustering
Big data

\textbf{ABSTRACT}

Free-floating micro-mobility as a mobility solution is becoming increasingly popular in cities. In this study, the travel patterns of free-floating electric bike-sharing service (FFEBSS) users before and during the COVID-19 pandemic were explored using big data and data mining. Existing real-time data studies provide a limited understanding of trip patterns and the characteristics of each user. Interpretations concerning the occurrence of life-changing events such as the COVID-19 pandemic are important. This study aimed to understand each user over 13 months comprising multiple time frames of market trends, seasonal change, and the COVID-19 pandemic outbreak. Multiple features were extracted from each user to explain the hidden data characteristics, and a data mining method was employed for clustering and evaluating user similarities with the extracted features. The results showed that FFEBSS users demonstrated a moderately stable travel pattern despite the COVID-19 pandemic, indicating the possibility of micro-mobilities being well adopted as our future urban transportation.

1. Introduction

The outbreak of the COVID-19 pandemic resulted in a devastating impact on global societies, economies, and urban systems. Furthermore, it gave rise to unprecedented scenarios. To mitigate the immediate impact of such an event and to understand the resulting human behaviors, evaluating the reaction of people to witnessing such unique situations is crucial. With the impact of the COVID-19 pandemic on urban mobility laid bare, this study explored the use of a free-floating electric bike-sharing service (FFEBSS) in Seoul, South Korea, before and after the COVID-19 pandemic outbreak. This study aimed to identify changes in the patterns of travel types, considering features including the spatial and temporal characteristics of individual user travels.

The present study investigates how the COVID-19 pandemic influences the travel patterns of FFEBSS users. An attempt was made to define behavioral patterns using big data, such as the global-positioning-system (GPS) coordinates of each ride (start and end points and path) and tracking of repetitive users for the past 13 months in Seoul, Korea. By using cluster analysis, groups not explicitly labeled in the data are segregated by similar traits to allow identification of unknown groups or confirmation of existing groups. Patterns from such groups allow for discoveries in the long-term perspective in micro-mobility adaptability and the practical and time-saving aspects as a mode for our future cities.

The methodological flow of this study follows the Knowledge Discovery in Databases (KDD) process by extracting useful information from databases (Fayyad & Stolorz, 1997). As can be seen in Fig. 1, the KDD process follows five basic steps. The first step, data selection, selects data that is valid and relevant to the study, in this case, the target users of the five service areas of Seoul are selected. Then, the target data goes through data cleaning and preprocessing, where the data is divided into eight different timeframes. The third step is data transformation where user-based behavioral features are extracted. Next, data mining involves the application of various statistical methods (Dehning et al., 2016), and in this study, a clustering methodology is employed. Finally, new knowledge is generated by interpreting and evaluating various patterns of the results.

The novelty of this study lies in the user-oriented analysis and time frame division. Distinct from most trip-based analyses, this study facilitates observation of travel patterns and characteristics of individual users based on their cumulative rides. The datasets are also divided into time frames of significance in consideration of the market cycle, seasonal change, and the COVID-19 pandemic. The such distinction
provides information on the impact of specific events. By presenting an understanding of the human travel patterns during the COVID-19 pandemic, this study can be useful for the future planning of regulations and urban infrastructure development as well as for preparing for unknown future events. However, as the data collected is limited to a specific service and location, the specifics of the results cannot be generalized.

The remainder of this paper is organized as follows. The following section discusses existing literature on the impact of the COVID-19 pandemic on travel behavior, FFEBSS, and FFEBSS patterns during the COVID-19 pandemic. Section 3 presents the overall methodology of the study. Following, the results including explanatory application from the cluster-based and time-based viewpoints are presented. The conclusion of the findings is discussed in the final section.

2. Literature review

2.1. Impact of COVID-19 pandemic on travel behavior

Several studies have been conducted on the impact of the environment on travel behavior, such as the impact of climate change on urban transport (Banister, 2011), the impact of weather conditions on active-transportation travel behavior (Saneinejad et al., 2012), and impact of weather on cycling and walking (Zhao et al., 2019). However, such studies may not be able to provide significant insights into the impact of the COVID-19 pandemic on travel patterns as the situation is unique owing to its resulting life-threatening complications, immediate implications, and fear of the unknown. The COVID-19 pandemic resulted in changes to the day-to-day lives of people, businesses, and to global trade and movements (Haleem et al., 2020). Thus, it interrupted the structure of regular life. Based on the concept of habit discontinuity, studies have shown that context change has the potential to make behavior-relevant information more salient and influential, which may lead to new choices and decisions (Müggenburg et al., 2015). For instance, studies have found that the COVID-19 pandemic has greatly affected low-income groups in terms of work, entertainment, transport, and many aspects of life due to loss of income leading to long-term effects, including limitation of travel choices and mobility (Brooks et al., 2020; De Vos, 2020; Jiang et al., 2017; Yang et al., 2021). Therefore, it is imperative to understand how people react in these possibly life-changing situations and find new ways to improve our communities in our plans, policies, and decision makings to make our cities more resilient (Fu & Zhai, 2021).

The first confirmed case of the COVID-19 pandemic in South Korea was reported on January 20th, 2020, and January 31st, 2020, for the city of Seoul. Since then, a large-scale outbreak emerged on February 20th, 2020, and immediate measures were implemented including distancing from physical contact with others and wearing face masks. No lockdowns proceeded but opening of schools was delayed and proceedings of flexible working arrangements, including work-from-home (WFH) systems, increasingly took place. In addition, South Korea released information on the traveled locations of those who tested positive to the public, and such data influenced people’s travel behaviors (Kim & Castro, 2020). By March 22nd, 2020, Social Distancing was strengthened, limiting group gatherings such as church services, entertainment centers, and gym facilities. As the number of cases decreased, the infection prevention act switched over to distancing in everyday life. Even so, the constant changes in the environment, people’s self-regulation, and governmental measures to restrict travel and social contact and to “flatten the curve” brought on behavioral changes in human travels well-worthy of discovery (Shakibaei et al., 2021).

Due to such changes in society, overall travel decreased. However, the patterns varied among modes. According to research, the usage of public transportation in Seoul, after the outbreak of the COVID-19 pandemic, decreased by approximately 32 % (Seoul Institute of Technology, 2020). Specifically, subway users in Seoul during the year 2020 decreased by 27.4 % compared to that in 2019. On the other hand, as for shared mobilities, travels increased by approximately 24 %. As for personal mobility (PM) services, the number of active users (MAUs) by major electric kickboard sharing service companies increased in the year 2020 compared to the previous year. Although there was a decreasing pattern from February 2020 (after the first outbreak of the COVID-19 pandemic), usage in April soared to its highest. Another study has revealed that after the COVID-19 lockdown, people are changing their habits in favor of a healthier and more sustainable lifestyle (Bergantino et al., 2021). The results also provide a possible estimation that PM
services are less influenced by external events such as the COVID-19 pandemic, rapidly rising as a sustainable transportation mode.

2.2. Free-floating electric bike-sharing service (FFEBSS)

FFEBSS, the type of service analyzed in this study, is an emerging personal mobility service that is experiencing rapid growth owing to its positive effects on mobility, the environment, and health (Bocker et al., 2020; Shaheen et al., 2012). In addition to being regarded as an economical, flexible, convenient, and sustainable travel mode, bike-sharing allows socially distant travel while demonstrating potential in mitigating problems such as air pollution and traffic congestion and in supporting multimodal transport connections (Jobe & Griffin, 2021; Ma et al., 2020). Regarding bike-sharing usage on health, cycling entails physical activity, and such a benefit plays an important role in the transition to more sustainable and healthy transport habits (Nikiforiadis et al., 2021). Bike-sharing systems can also increase the resilience of the transport system, maintaining its function in response to external impacts (Mattsson & Jenelius, 2015).

2.3. FFEBSS patterns during the COVID-19 pandemic

Discovering factors such as type of travel and travel purposes are not easy to define despite their importance and less has been covered on bike-share systems as sustainable transportation means (Kutela et al., 2021). Such characteristics also act in part to expect fast growth even during the COVID-19 pandemic and contribute to the resiliency of public transportation systems (Christoforou et al., 2021; Yan et al., 2021).

According to a case study done by New York's Citi Bike, the bike-sharing service proved to be more resilient than the subway system (Teixeira & Lopes, 2020). The COVID-19 pandemic has caused virus-wary travelers to stay away from public transit and FFEBSS can act as an option to prevent conversion into using a private car or a taxi (Liu et al., 2020; Nikiforiadis et al., 2020; Yan et al., 2021).

The data collected from free-floating bike-sharing (also known as dock-less bike-sharing) services is also characterized to be extensive, accurate, and informative, thus providing a user-friendly service and a method of investigating user characteristics and behaviors (Xing et al., 2020). Bikes also have the potential to solve the “last mile” problem, in that they can replace cars on short-distance trips (Flamm & Rivasplata, 2014; Yu et al., 2021).

Several studies have reported that many users of bike-sharing systems use them for leisure and not for commuting, which suggests that bike-sharing systems promoted trips that would not have been made in the absence of such services (Flamm & Rivasplata, 2014; Noland et al., 2011; Wang & Zhou, 2017). However, there are opposing opinions that, while shared-scooter usage patterns suggest usage for leisure, recreation, and tourism activities, bike-sharing is primarily used for commuting (McKenzie, 2019).

3. Methodology

3.1. Data selection and pre-processing: time-frame division

The data used in this study were collected from an FFEBSS company, Elecle (https://elecle.bike/). The obtained dataset comprised trips made in Seoul from April 5th, 2019, to May 18th, 2020, i.e., approximately 13 months. The provided data included information on the riding identification (ID), user ID, age, gender, start time, end time, minutes traveled, traveled distance, start area name, start location, and end location. The riding ID represents each tracked trip, and the user ID represents tracked repeat users. The entire dataset includes 37,046 free-floating trips. However, data considered as possible errors or inappropriate for this study were excluded to obtain a better analysis. The possible errors included trip data with zero duration or zero travel distance and travel shorter than the shortest distance from the start location to the end location. Such data were only observed during the beginning, and they can be attributed to glitches that may have occurred in the launching of a new service. The shortest distance comparison was calculated using a Euclidean distance metric, which is explained in subsequent sections. The data deemed inappropriate included single-time users. Single-time riders comprised users who only used the service once. Such users were excluded from the analysis as they were considered inappropriate for explaining changes in travel behavior. A final sample of 34,888 trips and 1642 users was used.

The FFEBSS areas in Seoul (the capital and largest metropolis of South Korea) comprise five districts: Mapo, Seodaemun, Yongsan, Jongno, and Jung-gu, as shown in Fig. 2. Seoul has a population of 9.7 million (2020) with an area of 605 km², and the five districts have a total population of 1.25 million (2019). It must be taken into consideration that the number of service users is not necessarily equal to the living population within the service areas as they may have traveled from outer districts or cities. The right side of the map in Fig. 2 shows subway stations and lines within Seoul in addition to the service areas.

On examining the social demographics of the users, their ages were found to be between 15 and 89 years with a mean of 29 years; 68 % of them were in their 20s, while 21 % were in their 30s. In terms of gender, 67 % of the users were male. These users were categorized into eight timeframes by trip date to appropriately compare the differences in behaviors. The timeframe division was made as described in Table 1, each based on a combination of service market cycle, seasons, and the COVID-19 pandemic. Along with considering such factors, the total count of days for each timeframe was as evenly distributed as possible. With the division, a single user can make trips at different timeframes, therefore the user count of the full dataset does not equal the sum of the users of each time frame. The time frame could be largely divided into two phases with $t_1$–$t_5$ being before the COVID-19 pandemic and $t_6$–$t_8$ during. As can be observed in the cumulative graph of trips in Fig. 2, a slow increase is observed in $t_1$–$t_3$ as the service is introduced to society. During $t_4$, the fall season, a significant increase is observed and the number of trips per user per week also appears noticeable; however, after $t_5$, $t_6$ and $t_7$ show a slow increase as the weather becomes cold and the COVID-19 pandemic occurs.

3.2. Data transformation: behavioral feature extraction

Several behavioral features were extracted from the user data in a manner that best expresses the hidden characteristics of the given data. In this study, a total of 11 features were taken into consideration: duration, distance traveled, speed, directness, frequency of use, peak hours, day of the week, first-mile and last-mile (FMLM), and activity space (area and aspect ratio). The combination of these features allows the use of data-oriented assumptions regarding concepts such as the purpose of the trip, human spatial behaviors, and repetitive-use patterns.

3.2.1. Duration

“used_min” represents the average travel time (how long the trip lasted) of the user in minutes.

3.2.2. Distance travelled

“distance(avg)” is the average distance traveled by the user in kilometers.

3.2.3. Speed

“speed” is calculated using the “used_min” and “distance(avg)” per trip and then averaged for each user (expressed in kilometers per hour).

3.2.4. Directness

“dtoeuc” is calculated by comparing the shortest distance between the start and end locations with the Euclidean metric (length of the line segment connecting two points) to the actual distance traveled. The values per trip are then averaged for each user. These values range from...
0 to 1. When this value approaches 1, it indicates that the trip was almost equivalent to the shortest path. The dtoeuc feature can be used to determine how efficiently one travels. More efficient travel usually indicates a stronger purpose of the trip, and less-efficient travel can be attributed to flexible trips such as trips that are made for the purpose of travel itself.

3.2.5. Frequency of use
“Frequency” indicates the average number of trips per day. However, it is different from how many trips a user made in a day. The purpose of this feature is to determine how often the users used the service during each time frame. If a user made 10 trips in 10 days, the frequency value would be 1, but this could also mean that they made five trips in 1 day and none on another day. Thus, the trip counts per user during the time frame were divided by the days in the time frame under consideration. This feature can have values as low as zero to those as high as infinity.

3.2.6. Peak hours
“peak (%)” represents the travels during peak hours per user. Peak hours comprise the busiest hours of the day during which most traveling occurs via automobiles, public transportation, or other possible modes. It is also usually the time for commutes. The peak hours used in this study are from 7:00 to 9:00 and 17:00 to 20:00. This feature is expressed in percentages and helps to identify commuters. Those traveling during peak hours are more likely to be traveling to work or school, which are activities that are repetitive and least affected by an exterior event.

3.2.7. Day of week
“weekend (%)” represents the amount traveled on weekends per user. This feature is also expressed in percentages, and a higher value indicates a greater number of trips made on weekends while a lower value indicates that more trips were made on weekdays. The day of the week is also a crucial factor affecting travel behavior, in that trip activities and purposes on weekdays and weekends differ significantly. The majority of trips made on weekends can be defined to be having the purpose of leisure (Zhong et al., 2008).

3.2.8. FMLM
The first and last miles respectively comprise the beginning and end of an individual trip that is usually connected with public

---

**Table 1**

| Time Frame   | Season              | Period                | Days | Trips | Users | Trips per User (Average per week) |
|--------------|---------------------|-----------------------|------|-------|------|-----------------------------------|
| t1 Service Launch | Spring             | 04/05/2019–05/19/2019 | 45   | 1025  | 137  | 1.2                               |
| t2 Expansion  | Spring and Summer   | 05/20/2019–07/19/2019 | 61   | 2472  | 278  | 1.0                               |
| t3 Growth     | Summer and Fall     | 07/20/2019–09/19/2019 | 64   | 3318  | 450  | 0.8                               |
| t4 Exuberance | Fall                | 09/20/2019–11/19/2019 | 61   | 11,225| 1037 | 1.2                               |
| t5 Contraction| Fall and Winter     | 11/20/2019–01/19/2020 | 61   | 7222  | 848  | 1.0                               |
| t6 COVID-19 Outbreak| Winter          | 01/20/2020–02/19/2020 | 31   | 2104  | 454  | 1.0                               |
| t7 Super Spread| Winter and Spring   | 02/20/2020–03/21/2020 | 31   | 2101  | 466  | 1.0                               |
| t8 Social Distancing| Spring         | 03/22/2020–05/18/2020 | 58   | 5421  | 832  | 0.8                               |
| Full Dataset  |                     | 04/05/2019–05/18/2020 | 412  | 34,888| 1642 | 0.4                               |

---

**Fig. 2.** Map of the FFEBSS areas - five districts within Seoul
Upper Left: A broader view of the location of the districts within Seoul (Map of Seoul, South Korea. kakaomap, 2020, map.kakao.com).
transportation. The first mile can be defined as the initial point of the
trip to a public transportation station and the last mile can be defined as
the station to the destination. In this study, the definition of “arriving at”
or “departing from” a subway station has been applied to when the start
or end location, respectively, is within 180 m from a station, which is the
weighted average distance between stations in a transfer hub in the
metropolitan area with an LOS C (standard developed by Ministry of
Land, Infrastructure and Transport of South Korea). As the subway sta-
tion point coordinates used are based on a representative location
among multiple exits of each station, the 180 m presumption considers
the most possible exit. Fig. 3 represents an example of Hongik Univ.
station point location where the station is known to have the longest
distance transfer within Seoul. Fig. 4 (c) shows an example of a first-mile
travel path of a user considered to have transferred from FFEBS to
getting on the subway at Hongik Univ. station.

“firstmile (%)” is calculated as the number of trips made from a start
point to a subway station per user, and “lastmile (%)” represents the
number of trips made from a subway station to an endpoint per user.
These features can be used to define trips that did not comprise the
entire journey of the user. Fig. 5 presents each start location and end
location of the entire dataset in blue dots, and all the subway stations
within a radius of 180 m from them in red circles. The feature values
would comprise the proportion of blue dots within the red circles per
user.

3.2.9. Activity space

The activity space is usually defined as the area within which people
travel for their daily activities. The shape of the area, its definition, and
its interpretation vary. In this case, a confidence ellipse (also known as
as a standard deviation ellipse) is used to measure the activity space of each
user according to their travel locations. The shape of the ellipse is
determined using a Pearson correlation coefficient (Schelp, 2018,
which is presented in Eq. (1).

\[ \rho = \frac{\text{cov}_{xy}}{\sigma_x \sigma_y} \]  

(1)

where cov_{xy} is the covariance, which is divided by the product of the
standard deviations of the variables (\( \sigma_x, \sigma_y \)) to yield the Pearson corre-
lation coefficient (\( \rho \)). The axes of the ellipse are then determined based
on this value, as shown in Eqs. (2) and (3):

\[ \text{Horizontal radius} = \sqrt{1 - \rho} \]  

(2)

\[ \text{Vertical radius} = \sqrt{1 - \rho^2} \]  

(3)

The size of the ellipse would then be determined based on the desired
standard deviations. In this study, both the start and end points of trips
were plotted in the calculation used to obtain the ellipse to better define
each user’s activity space, and two standard deviations (95 % of the
data) were selected as the standard for drawing the ellipse.

Using this ellipse, two features were extracted: area and aspect ratio.
“area” represents the area of the drawn ellipse per user. This feature
describes the activity space of the user, where small values indicate a
small boundary of the activity travel and large values indicate a large
boundary of the activity travel. As an, the further one travels, the larger
the area. From the example shown in Fig. 6, it can be observed that User
15 (left) has a smaller boundary of activity travel compared to User 375
(right). Aspect ratio, “h/w”, represents the activity-space shape, i.e., the
ratio of the length of the minor axis to that of the major axis, and it is
used to define the activity pattern. The value of this feature ranges from
0 to 1. A value closer to 1 indicates a shape closer to a perfect circle, i.e.,
it does not comprise a distinctive travel pattern. In contrast, a value
closer to 0 indicates a shape closer to a line, which indicates repetitive
travel in a space with a specific purpose, i.e., the user regularly travels
from similar start and end locations. For example, a user who uses the
service as a means of commute would most likely start or end the travel
at their home or workplace. In Fig. 6, User 15 (left) has a “h/w” value
closer to 1 and User 375 (right) with a value lower than 0.5. It can be
indicated that User 375 has a more repetitive pattern in travel location.
These features thus facilitate interpretations of trip behaviors based on
how large each user’s activity boundary is and whether their activities
are repetitive.

3.3. Data mining: K-means clustering

In data mining, a clustering method is used to identify patterns to
provide insight into the basic structure of the data (Aytac, 2020;
Govender & Sivakumar, 2020; Mohri et al., 2021). Generally, clustering
methods can be categorized as flat and hierarchical algorithms (Jafar-
zadegan et al., 2019). The K-means algorithm is the simplest and most
commonly used algorithm that repetitively assigns patterns to clusters
based on the similarity between the pattern and the cluster centers until
a convergence criterion is met (Jain et al., 1999). This method provides
the advantages of low complexity, fast computation, ability to handle
large datasets, and adjustability of the cluster membership (Govender
& Sivakumar, 2020). In this study, extracted features per timeframe per

![Fig. 3. Cumulative daily trip count from April 5th, 2019, to May 18th, 2020 (Data provided by Elecle).](image-url)
user were clustered to similar user groups using the K-means algorithm. As the features were divided into each timeframe, although all 13 month data were input at once, the same users of different timeframes were considered different. However, as each user is labeled, the individuals can be tracked after the clustering process.

Before the clustering analysis of the features, additional data transformations were made using correlation analysis, scaling, and normalization to improve data integrity. Features with strong correlations, that do not convey additional information, were excluded. Among the several scaling methodologies a logarithmic scaling method was used for solving the problem of skewness towards larger values for “distance(avg).” Using the logarithmic scale, each interval of the data is

Fig. 4. FMLM example at Hongik Univ. station point location (Map of Hongik Univ. station, South Korea. kakaomap, 2020, map.kakao.com)
(a). Example location – Hongik University Station, Seoul
(b). 180 m radius of example location
(c). Example of user travel information map.

Fig. 5. FMLM plot: first-mile plot (left), last-mile plot (right) (Map of Seodaemun-gu, South Korea. kakaomap, 2020, map.kakao.com).
increased by a factor of the base of the logarithm. In this case, a base of ten was used, thus making the data easier to comprehend. As the K-means algorithm uses Euclidean distance, which is highly prone to irregularities in the size of various features (Patel & Mehta, 2011), the min-max normalization method was applied.

3.3.1. K-means

In the following section, a K-means clustering method was used to identify the similarities between the users grouped by their features. The K-means clustering process works as follows. k number of clusters is pre-specified by the researcher. Then, k random data points are initialized as centroids. The sum of all the distances between the data points and centroids is computed, and each data point is assigned to its closest centroid, thus forming a cluster. The centroids are re-centered using the average of all the data points belonging to each cluster, and the entire process is repeated. The iteration continues until no changes are made to the clusters.

The main objective of the K-means algorithm is to obtain the smallest variance difference between the Euclidean distance of each cluster to optimize the function in Eq. (5):

$$\sum_{i=1}^{k} \sum_{x \in S_i} ||x - \mu_i||^2 \quad (5)$$

where n given data points \(\{x_1, x_2, \ldots, x_n\}\) are assigned to k clusters \(S = \{S_1, \ldots, S_k\}\), and \(\mu_i\) represents the mean of the data points in \(S_i\).

This algorithm has several drawbacks. The first drawback is the need to pre-determine the number of clusters. Depending on the k number of clusters set, the result may vary in terms of how well the similarities are grouped. Furthermore, as this algorithm makes use of Euclidean-distance-based measurements to determine similarities, preprocessing becomes crucial to ensure that the datasets considered have equal measurement units. Lastly, the random initialization of centroids may result in a local minimum instead of convergence at a global minimum.

Several methods can be used to overcome these problems and obtain the optimal clustering result. For selecting a suitable number of clusters, two methods are used in this study: the silhouette coefficient method and the elbow method. A silhouette analysis presents the separation distance between the resulting clusters, wherein a silhouette coefficient value near +1 indicates a large distance from the neighboring clusters and is thus a desirable value. For better precision, the elbow method was also applied, which comprises the use of the sum of the squared distance (SSE) between the data points and assigned clusters. The point in the graph at which the SSE starts to flatten out provides the optimal number, 6, that we require. The unit normalization problem of the features was already considered during the preprocessing. For the initialization of the centroids, the method presented in this study was used to provide several different initializations of the centroids and for obtaining the result with the lowest sum of the squared distance.

4. Results

4.1. User-based behavioral features

Fig. 7 presents the mean value plot of the behavioral features. It can be observed that the values demonstrate an apparent impact of the COVID-19 pandemic. For example, “Duration (min)” exhibited the lowest minutes (average of 10 min) during the COVID-19 outbreak in Korea (t6), in addition to the “Weekend (%)” showing a significant decrease during this period. “Area” and “Aspect Ratio” also showed significant change after the outbreak (t6) and moving on to social distancing (t8). “dteoue” exhibited constantly high values during the COVID-19 pandemic, revealing that direct destination trips were made. “firstmile (% )” and “lastmile (% )” showed constantly high values (25–27 %) during the pandemic. In particular, “lastmile (% )” demonstrated the highest values during the pandemic as compared to all the other time frames.

The patterns also show relative observations of the weather. In Seoul, the year divides into four distinct seasons: hot and humid summers, cold and snowy but dry winters, and spring and autumn in between. The temperature during the duration of the study varied from -9.5 °C to 29.9 °C and precipitation up to 63.2 mm and the total user count showed a related pattern to such weather. Features such as “duration” and “weekend” also showed correlations with the weather. The duration of the trips appeared longer during warmer weather and weekend averages decreased in both the hottest and coldest temperatures. Appendix A and Appendix B provide more descriptive statistics of the extracted behavioral features.

4.2. User classifications and their travel patterns

Before performing K-means clustering, two features (“used_min” and “area”) were eliminated due to a strong correlation with other features. A total of nine features were selected and scaled for analysis. In pre-determining the k number of clusters, a k value between 3 and 8 was selected to obtain the optimal value while using both the silhouette
scores and the elbow method. The elbow method presented flattening between 6 and 7 and the silhouette score was also shown to have its highest value at k = 6 (0.2416). Therefore, six was selected as the number of clusters. In consideration of the need for different initializations of centroids, 10 random initial values were applied with maximum iterations up to 300. The result with the lowest sum of squared distance was selected for ease of comprehension.

Fig. 8 presents the results of the clustering process. The bar plots were drawn based on the cluster centers of each feature. The mean values of the behavioral features in each of the six clusters are provided in Appendix C. On observing each cluster and their comparison, the results thus obtained were found to be quite distinctive. From the observed characteristics, the cluster groups were each named “Community Travelers,” “Commuters,” “First Milers,” “Last Milers,” “Weekenders (FMLM),” and “Weekenders.”

The first four and last two clusters can be separated based on the day of the week. The two weekend user groups can then be distinguished based on the use of public transportation. “Weekenders (FMLM)” exhibit high first-mile and last-mile portions and low speed, while “Weekenders” exhibit low first-mile and last-mile portions and an average speed. “Weekenders (FMLM)” can be interpreted as weekend leisure travelers who use the subway as a main mode of transport. In contrast, “Weekenders” can be understood as weekend leisure travelers who use the FFEBSS itself as a mode of transport to a specific destination, or the trip itself as a purpose. When a trip itself is defined as the purpose, the case comprises cycling as the main activity. Both groups are defined as leisure travelers not just because of the day of the week but also because these users exhibit low “h/w,” “dtoeuc,” and “frequency” and high “distance(avg)” values. Such behavior can be attributed to trips where the origin and destination are different and non-frequent, but the travel distance is also not traveled using the shortest-path behavior.

The groups that traveled during weekdays can be divided into two groups comprising those that use public transportation and those that do not. “First Milers” and “Last Milers” can be easily identified by their features. In addition to having high “firstmile (%),” “lastmile (%),” and “h/w” values, these groups also exhibit high “peak ( %) but low “frequency” values. The low “frequency” factor shows that these users should not be classified as commuters. Moreover, the activity space characteristics of these users exhibit low “h/w” travels but with high “dtoeuc,” which indicates that the non-frequent, destination trips are made such that the traveled path is close to the shortest path.

“Community Travelers” are named based on multiple features: high “h/w,” “dtoeuc,” “speed,” “frequency,” “peak (%)” and low “firstmile (%)” and “lastmile (%).” High “h/w” and “dtoeuc” travelers can be defined as those who travel highly efficiently within a boundary. People who often travel within a community, during peak hours, and at a high speed can be considered workers who live and work within a similar activity-space area. They can also be interpreted as students who use the FFEBSS within their campus. In contrast, “Commuters” exhibit a low “h/w” and high “dtoeuc” during peak hours. This pattern appears to be very similar to that of the first mile and last mile travelers, in that the trips are destination specified and highly efficient in terms of the path. However, these users exhibit low “firstmile (%)” and “lastmile (%)” percentages, which indicates that the FFEBSS is used as a mode of transport by these commuters. On observing the users, the clusters did not exhibit distinctiveness in social demographics.

Based on these findings, the percentage change over time of the cluster groups was observed (Fig. 9). Such observations facilitated the identification of travel patterns caused by a combination of seasonal comparisons, market trends, and the COVID-19 pandemic outbreak. The most dramatic changes occurred during t4 (exuberance, fall season) and during the COVID-19 pandemic (t6–t8). t4 exhibited a sudden increase and then decrease in the percentages of “Community Travelers” and “Weekenders (FMLM),”

From the point when the COVID-19 pandemic outbreak was announced in Korea (t6), multiple patterns were observed. “Community Travelers”, “Commuters”, and “Weekenders” demonstrated the most noticeable changes. In observing “Community Travelers” during the COVID-19 pandemic outbreak, the largest decrease in percentages was observed over the entire eight timeframes under consideration. Compared to an average of 38.6 % from Service Launch to Contraction, “Community Travelers” exhibited a cluster percentage of 28.6 % and remained at the lowest points throughout the COVID-19 pandemic. On the other hand, “Commuters” showed their highest values during the outbreak. With a percentage of 29.7 %, it was the only time having a greater percentage than “Community Travelers”, although
decreases were observed afterward, the percentages stayed within the fluctuation range before the COVID-19 pandemic. “Weekenders” showed the highest percentages during Super Spread. While the average during the rest of the time frames was 11.4 %, during t7 showed 18.5 %, going above the fluctuation range.

Looking further into the data, the social demographic features of the users were observed. Fig. 10 presents the gender distribution of clusters for each time frame where the percentage sum of all cluster groups for each gender per timeframe is 100 %. Noticeable patterns were observed among “Community Travelers” and “Commuters”. Although both genders showed similar patterns throughout the time frame, the rate of change differed especially from t5 to t6. Of the drop in percentages during the COVID-19 pandemic outbreak, female “Community Travelers” showed a larger drop of 12.8 % (34.9 % to 22.1 %) while male “Community Travelers” showed 8.2 % (38.5 % to 30.4 %). The male “Community Travelers” also showed greater recovery of percentages throughout the COVID-19 pandemic. Female “Commuters” during Super Spread showed a value as low as 22.12 % while the average of the rest of the time frames was 30.32 %. Male users on the other hand exhibited an 8.61 % increase during the COVID-19 pandemic outbreak. Both genders appeared to return to average during the Social Distancing period.

The age distribution was analyzed under the age of 40 (Fig. 11). The users of age 40 and up were excluded since the sample size was too small to properly define a pattern. In observing the “Community Travelers”, users in their 30s exhibited the smallest rate of change during the COVID-19 pandemic outbreak (3.43 %), but the largest rate of change during the Super Spread (8.41 %). The largest decrease in percentages during the outbreak appeared among the age of 25 to 29, dropping from 40.2 % to 27.92. However, while all three age groups showed a drop in percentages during the outbreak, users in their 30s were the ones showing fluctuation in patterns. Users in their 10s and 20s showed stable percentages throughout the COVID-19 pandemic.

Contrary to the pattern of “Community Travelers”, “Commuters” in their 30s had the smallest fluctuation range during the COVID-19 pandemic. With a fluctuation range of 8 % throughout all the time frames, users in their 30s showed an even smaller fluctuation range of 2.18 % during the COVID-19 pandemic. On the other hand, age groups of 15 to 24 showed a fluctuation range of 16.6 % throughout the entire analysis period and 9.13 % during the COVID-19 pandemic, and age groups 25 to 29 showed 14 % and 10.69 % ranges. Therefore, commuters in their 30s could be observed as the most stable group with or without the influence of the COVID-19 pandemic.

The age group distribution of the “Weekenders” also exhibited a noticeable result. While users in their 10s and 20s showed a drop in percentages during the COVID-19 pandemic outbreak, those in their 30s had an increasing pattern.

In further analyzing patterns during the COVID-19 pandemic, Fig. 12 presents the number of new riders and the difference between new and non-riders. The non-riders are those who used the service in the previous time frame but not in the next time frame. After the outbreak of the COVID-19 pandemic, “Weekenders” show a noticeable number of new riders compared to non-riders (t6 → t7). “Community Travelers” and
“Commuters” revealed a significantly higher number of non-riders from Contraction (t5) to the outbreak of the COVID-19 pandemic (t6), but the opposite from Super Spread (t7) to Social Distancing (t8). Additionally, in comparing the overall new and non-users before and after the outbreak of the COVID-19 pandemic, 36% among the entire 13 month users stopped riding after the outbreak of the COVID-19 pandemic, while 20% were entirely new users that appeared after the outbreak. Considering that the data provided consisted of more days before the COVID-19 pandemic than after, the percentage difference between new and non-users do not seem significant.

A noteworthy observation is that while changes can be identified and interpreted, the overall pattern in the cluster groups remains consistent. Other than the fact that the quantity of FFEBSS users has changed most significantly during the COVID-19 pandemic, the difference between the change rate of percentages before and during the COVID-19 pandemic does not seem significant.

5. Conclusion and discussion

5.1. Conclusion

This study was conducted to obtain information regarding changes in the travel patterns of FFEBSS users owing to the impact of the COVID-19 pandemic. To be noted in this analysis is the focus on users, not just the individual trips, along with the extraction of behavioral features that allow for the investigation of more specific findings during specific time frames. A total of 1642 users (34,888 trips) within five service areas of Seoul, Korea, during eight different time frames (t1–t8) were clustered using nine user-based behavioral features: distance, speed, frequency, peak hours, day of the week, first mile, last mile, directness (dtoeuc), and activity-space aspect ratio (h/w). The cluster results were then analyzed using cluster-based and time-based analyses.

On observing the behavioral feature descriptions by time frames, “peak (%)” exhibited the lowest value during the first two months of the COVID-19 pandemic with a particularly noticeable decrease during the super spread of the virus. This shows that an unprecedented situation can affect and change the pattern of known stable features such as peak hour travels. Moreover, the recovery of “peak (%)” in the third month of the COVID-19 pandemic indicated the duration of the impact on peak travel for FFEBSS users in Seoul. Another feature that showed a considerable change during the COVID-19 pandemic is the decline of the “area” values. Such behavior can be attributed to the impact of the spread of the unknown virus on users’ willingness to travel farther in larger activity boundaries. “dtoeuc”, “firstmile (%),” and “lastmile (%)” exhibited constantly high values during the COVID-19 pandemic, indicating positive attitudes towards FFEBSS in the use of destination-focused and FMLM trips.

Observing the six different travel behavior clusters, resulting from the application of the K-means algorithm, suggest the overall stability of FFEBSS user patterns despite the COVID-19 pandemic outbreak. Change in activity spaces, stability in commuting use for age groups in their 30s, first and last-mile-purpose new rider increases, and fast recoveries were observed. However, the most significant changes were observed in terms of quantity. On comparing the ride counts of the data used in this study to the mobility trend reports of Seoul (Apple, n.d.) and the subway use data of Seoul provided by the Seoul Metro, the travel flow during the COVID-19 pandemic was found to be very similar. This indicates that the changes in the use of the FFEBSS were not due to the specific characteristics of the service but due to the decrease in the amount of movement after the COVID-19 pandemic outbreak. In contrast, the travel patterns that define the behaviors of the service users demonstrated stability, especially in the case of commuters. Opposite to “Community Travelers” highly influenced by the COVID-19 pandemic, the pattern of “Commuters” showed that they are less impacted and inelastic to such exterior conditions.

5.2. Discussion

In an effort to avoid contracting the COVID-19 pandemic, people demonstrated a reluctance towards ridesharing for sanitary reasons and have changed their mobility patterns significantly. A survey conducted with respondents across seven countries has shown that less than 10% of the survey respondents believe that carsharing, ridesharing or public transport are safe after the COVID-19 pandemic outbreak. For example, ride-hailing companies in multiple geographies have experienced 60–70% declines in passengers during the COVID-19 crisis.
(McKinsey&Company, n.d.). Despite such uncertainty, the stability of user patterns in this study suggests that the FFEBSS has advantages that could allow it to become the next-generation mobility service. As policies and regulations for micro-mobility sharing services are piloted, refined, and implemented, this study indicates the potential of micro-mobility for becoming a critical component of the wider mobility spectrum. Moreover, the result of the most stable travelers, i.e., the commuters, exhibiting the highest stability in FFEBSS usage patterns suggests that a greater interest from governments and service providers is required for providing a more extensive, accurate, and user-friendly urban-mobility system in addition to providing a smart solution to the “last-mile” problem.

5.3. Limitations

While the proposed method has effectively demonstrated the travel-pattern findings in a data-oriented manner, it has some limitations. One limitation is that the collected data were obtained from a newly introduced service. This may have affected the data reliability for the first few months after the launch of the service. The scope of the analysis is another limitation. This study comprised the use of only five districts within Seoul and does not represent the entire population. This suggests that the results cannot be generalized, and additional research is required to be conducted with a longer-term dataset within a broader scope of the service area.

In extracting behavioral features, owing to a lack of data, the first and last mile travels were limited to be defined as travels to and from a subway station. There is a possibility of travel beginning or ending near a subway station. However, the definition applied considers the weighted average distance between stations in a transfer hub. In using the “transfer distance” as a reference, it is assumed that activities are not possible within the distance other than “transferring”. Defining “Activity

Fig. 10. Gender distribution by time frame.
"Space also contains limitations in that activity patterns may vary according to different urban locations. Another limitation of this study is its lack of comparative modes of transport, which can be remedied in future research. In a research project on mobility behavior in Switzerland (Molloy et al., 2020) that focused on the impacts of the coronavirus on mobility behavior, comparisons were made via multiple modes. The change in the average daily distance percentages per week was observed for six different modes. The results thus obtained showed significantly positive percentages for bicycle usage while all the other modes including travel by bus, car, and walking were below zero. On comparing the data used in this study with that of the MOBIS study, it was observed that the change in the daily distance of FFEBSS users was much above zero just as in the case of the bicycle users of the MOBIS study. Future research that comprises the use of a pool of participants that matches that in this study would be relevant for observing the impact of the COVID-19 pandemic on specific modes of transport and better understanding the resulting human travel patterns.

Fig. 11. Age distribution by time frame.

Fig. 12. New & non-riders during the COVID-19 pandemic.
CRediT authorship contribution statement

Seung Eun Choi: Conceptualization, Data curation; Formal analysis; Methodology; Visualization, Original Draft Dayoung Seo: Data curation Jinhee Kim: Conceptualization; Methodology; Project administration; Validation; Supervision; Review & Editing.

Declaration of competing interest

The authors certify that they have NO conflicts of interest and other potentially conflicting interests, including specific financial interests and relationships and affiliations relevant to the Safety Science (e.g., employment/affiliation, grants or funding, consultancies, honoraria, stock ownership or stock options, expert testimony, royalties, or patents filed, received, or pending).

Data availability

The authors do not have permission to share data.

Acknowledgments

This research was supported by the Transport and Logistic Research Project (project number: 18TLRP-B148676-02) funded by the Republic of Korea Ministry of Land, Infrastructure and Transport (MOLIT) and Korea Agency for Infrastructure Technology Advancement (KAIA).

Appendix A. Behavioral feature descriptions

| Period | Season | Time Frame | Spring | Summer | Fall | Winter | Spring |
|--------|--------|------------|--------|--------|------|--------|--------|
|        |        | t1 | t2 | t3 | t4 | t5 | t6 | t7 | t8 |
|        | Service Launch | Expansion | Growth | Exuberance | Contraction | COVID-19 Outbreak | Super Spread | Social Distancing |
| Temperature (°C) | 12.98 | 20.65 | 24.82 | 13.48 | 0.76 | 1.04 | 4.28 | 10.84 |
| Precipitation (mm) | 1.54 | 1.90 | 8.04 | 2.15 | 1.42 | 0.80 | 1.32 | 1.18 |
| FEATURES | Number of users | 137 | 278 | 450 | 1037 | 848 | 454 | 466 | 832 |
| Speed (km/h) | Mean | 12.82 | 14.69 | 15.47 | 14.44 | 15.17 | 15.74 | 15.39 | 13.53 |
| Min | 0.23 | 1.16 | 0.65 | 0.18 | 0.94 | 0.68 | 0.62 | 0.43 |
| Max | 47.90 | 30.60 | 38.96 | 37.97 | 33.37 | 31.59 | 56.26 | 30.87 |
| Duration (min) | Mean | 13.50 | 13.33 | 15.59 | 18.48 | 11.73 | 10.08 | 11.67 | 16.97 |
| Std | 19.01 | 21.19 | 39.49 | 136.66 | 46.72 | 14.35 | 19.18 | 38.42 |
| Min | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Max | 215.00 | 221.00 | 707.00 | 4357.00 | 1270.00 | 188.00 | 232.00 | 641.00 |
| Distance (km) | Mean | 1.83 | 2.27 | 3.35 | 5.67 | 1.20 | 1.48 | 1.98 | 2.74 |
| Std | 0.08 | 0.04 | 0.05 | 0.04 | 0.05 | 0.01 | 0.03 | 0.01 |
| Min | 0.00 | 14.16 | 16.89 | 40.77 | 151.17 | 13.03 | 24.52 | 27.04 |
| Max | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Temperature (°C) | Mean | 13.50 | 13.33 | 15.59 | 18.48 | 11.73 | 10.08 | 11.67 | 16.97 |
| Std | 19.01 | 21.19 | 39.49 | 136.66 | 46.72 | 14.35 | 19.18 | 38.42 |
| Min | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Max | 215.00 | 221.00 | 707.00 | 4357.00 | 1270.00 | 188.00 | 232.00 | 641.00 |
| Frequency | Mean | 0.17 | 0.15 | 0.15 | 0.14 | 0.13 | 0.14 | 0.16 | 0.17 |
| Std | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.03 | 0.03 |
| Min | 0.00 | 1.29 | 1.16 | 1.19 | 2.12 | 1.51 | 1.42 | 1.48 | 1.48 |
| Max | 0.00 | 14.28 | 13.57 | 13.79 | 14.56 | 12.17 | 11.45 | 10.39 | 13.11 |
| Peak (%) | Mean | 28.25 | 24.48 | 24.07 | 33.10 | 32.27 | 24.37 | 33.41 | 32.20 |
| Std | 34.11 | 31.54 | 31.37 | 34.16 | 34.91 | 32.96 | 38.77 | 36.09 |
| Min | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Max | 0.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| Firstmile (%) | Mean | 20.56 | 20.89 | 24.30 | 26.85 | 24.10 | 26.31 | 25.26 | 26.13 |
| Std | 26.66 | 30.47 | 31.69 | 29.88 | 31.22 | 35.24 | 34.64 | 33.01 |
| Min | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Max | 0.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| Lastmile (%) | Mean | 20.64 | 24.90 | 24.07 | 24.65 | 21.68 | 25.41 | 25.60 | 26.41 |
| Std | 30.13 | 33.38 | 32.19 | 30.19 | 30.16 | 33.86 | 34.98 | 33.55 |
| Min | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Max | 0.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| Area | Mean | 27.14 | 27.25 | 24.96 | 29.92 | 26.88 | 21.68 | 20.99 | 25.39 |
| Std | 19.12 | 19.87 | 19.88 | 18.16 | 19.19 | 19.60 | 19.73 | 19.49 |
| Min | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Max | 0.00 | 50.27 | 50.27 | 50.27 | 50.27 | 50.27 | 50.27 | 50.27 |
| Aspect Ratio (h/w) | Mean | 0.38 | 0.39 | 0.35 | 0.42 | 0.38 | 0.30 | 0.28 | 0.36 |
| Std | 0.31 | 0.33 | 0.32 | 0.31 | 0.31 | 0.32 | 0.33 | 0.32 |
| Min | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Max | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.99 | 1.00 |
Appendix B. One-way ANOVA of behavioral features

| Features          | Time Frame | Difference | P-value | Signif |
|-------------------|------------|------------|---------|--------|
| Speed (km/h)      | t1 vs t2  | 1.87       | 0.00    | **     |
|                   | t2 vs t3  | 0.78       | 0.35    |        |
|                   | t3 vs t4  | -1.03      | 0.00    | **     |
|                   | t4 vs t5  | 0.73       | 0.02    | *      |
|                   | t5 vs t6  | 0.57       | 0.41    |        |
|                   | t6 vs t7  | -0.35      | 0.95    |        |
|                   | t7 vs t8  | -1.86      | 0.00    |        |
|                   | t1 vs t2  | -0.17      | 1.00    |        |
|                   | t2 vs t3  | 2.26       | 1.00    |        |
|                   | t3 vs t4  | 2.89       | 1.00    |        |
|                   | t4 vs t5  | 1.03       | 0.00    | **     |
|                   | t5 vs t6  | 0.73       | 0.02    | *      |
|                   | t6 vs t7  | 0.54       | 0.41    |        |
|                   | t7 vs t8  | 0.13       | 1.00    |        |
|                   | t8 vs t9  | 1.03       | 0.00    | **     |
| Duration (min)    | t1 vs t2  | 0.17       | 1.00    |        |
|                   | t2 vs t3  | 0.17       | 1.00    |        |
|                   | t3 vs t4  | 0.22       | 1.00    |        |
|                   | t4 vs t5  | 0.01       | 1.00    |        |
|                   | t5 vs t6  | 0.01       | 1.00    |        |
|                   | t6 vs t7  | 0.01       | 1.00    |        |
|                   | t7 vs t8  | 0.01       | 1.00    |        |
|                   | t8 vs t9  | 0.01       | 1.00    |        |
| Distance (km)     | t1 vs t2  | 0.17       | 1.00    |        |
|                   | t2 vs t3  | 0.17       | 1.00    |        |
|                   | t3 vs t4  | 0.22       | 1.00    |        |
|                   | t4 vs t5  | 0.01       | 1.00    |        |
|                   | t5 vs t6  | 0.01       | 1.00    |        |
|                   | t6 vs t7  | 0.01       | 1.00    |        |
|                   | t7 vs t8  | 0.01       | 1.00    |        |
|                   | t8 vs t9  | 0.01       | 1.00    |        |
| Directness (dtoeuc)| t1 vs t2  | 0.03       | 0.53    |        |
|                  | t2 vs t3  | 0.02       | 0.65    |        |
|                  | t3 vs t4  | 0.05       | 0.00    | **     |
| Frequency         | t1 vs t2  | 0.04       | 0.00    | **     |
|                  | t2 vs t3  | 0.04       | 0.00    | **     |
|                  | t3 vs t4  | 0.06       | 0.00    | **     |
|                  | t4 vs t5  | 0.04       | 0.00    | **     |
|                  | t5 vs t6  | 0.04       | 0.53    |        |
|                  | t6 vs t7  | 0.04       | 0.53    |        |
|                  | t7 vs t8  | 0.04       | 0.53    |        |
|                  | t8 vs t9  | 0.04       | 0.53    |        |
| Peak (%)          | t1 vs t2  | 0.71       | 1.00    |        |
|                  | t2 vs t3  | 0.22       | 1.00    |        |
|                  | t3 vs t4  | 0.77       | 1.00    |        |
|                  | t4 vs t5  | 2.72       | 0.24    |        |
| Weekend (%)       | t1 vs t2  | 3.77       | 0.97    |        |
|                  | t2 vs t3  | 1.22       | 1.00    |        |
|                  | t3 vs t4  | 9.84       | 0.00    | ***    |
|                  | t4 vs t5  | 9.84       | 0.00    | **     |
|                  | t5 vs t6  | 9.04       | 0.00    | **     |
|                  | t6 vs t7  | 9.04       | 0.00    | **     |
|                  | t7 vs t8  | 9.04       | 0.00    | **     |
|                  | t8 vs t9  | 9.04       | 0.00    | **     |
| Firstmile (%)     | t1 vs t2  | 2.21       | 0.93    |        |
|                  | t2 vs t3  | 1.05       | 1.00    |        |
|                  | t3 vs t4  | 3.41       | 0.86    |        |
|                  | t4 vs t5  | 4.62       | 0.91    |        |
| Lastmile (%)      | t1 vs t2  | 0.87       | 1.00    |        |
|                  | t2 vs t3  | 0.87       | 1.00    |        |
|                  | t3 vs t4  | 0.87       | 1.00    |        |
|                  | t4 vs t5  | 0.87       | 1.00    |        |
|                  | t5 vs t6  | 0.87       | 1.00    |        |
|                  | t6 vs t7  | 0.87       | 1.00    |        |
|                  | t7 vs t8  | 0.87       | 1.00    |        |
|                  | t8 vs t9  | 0.87       | 1.00    |        |
| Area              | t1 vs t2  | 4.96       | 0.00    | **     |
|                  | t2 vs t3  | 4.96       | 0.00    | **     |
|                  | t3 vs t4  | 4.96       | 0.00    | **     |
|                  | t4 vs t5  | 4.96       | 0.00    | **     |
|                  | t5 vs t6  | 4.96       | 0.00    | **     |
|                  | t6 vs t7  | 4.96       | 0.00    | **     |
|                  | t7 vs t8  | 4.96       | 0.00    | **     |
|                  | t8 vs t9  | 4.96       | 0.00    | **     |
| Aspect Ratio (h/w)| t1 vs t2  | 0.04       | 0.71    |        |
|                  | t2 vs t3  | 0.04       | 0.71    |        |
|                  | t3 vs t4  | 0.04       | 0.71    |        |
|                  | t4 vs t5  | 0.04       | 0.71    |        |
Appendix C. Feature mean values of the six clusters

| Cluster            | Number of Users | Distance (km) | Speed (km/h) | Frequency | Peak (%) | Weekend (%) | Firstmile (%) | Lastmile (%) | Directroute (%) | Aspect Ratio (h/w) |
|--------------------|----------------|---------------|--------------|-----------|----------|-------------|--------------|-------------|-----------------|-------------------|
| Community Travelers| 1583           | 2.15          | 15.70        | 0.26      | 20.55    | 24.75       | 19.26        | 18.21        | 0.53            | 0.69              |
| Commuters          | 1092           | 2.25          | 14.40        | 0.10      | 9.94     | 7.62        | 3.88         | 5.95         | 0.52            | 0.15              |
| First Mileers      | 392            | 3.06          | 14.76        | 0.07      | 13.92    | 10.96       | 79.11        | 9.49         | 0.53            | 0.18              |
| Last Mileers       | 558            | 2.92          | 14.58        | 0.10      | 13.84    | 9.71        | 38.78        | 82.29        | 0.53            | 0.21              |
| Weekenders         | 561            | 2.70          | 13.54        | 0.05      | 2.00     | 90.26       | 3.59         | 22.08        | 0.48            | 0.18              |
| Weekenders (FMLM)  | 316            | 3.41          | 13.04        | 0.05      | 1.95     | 89.92       | 75.63        | 39.93        | 0.49            | 0.27              |

References

Apple. (n.d.). COVID-19—Mobility Trends Reports—Apple. Retrieved August 28, 2020, from https://covid19.apple.com/mobility.

Aytaç, E. (2020). Unsupervised learning approach in defining the similarity of catchments: Hydrological response unit based k-means clustering, a demonstration on Western Black Sea region of Turkey. International Soil and Water Conservation Research, 8(3), 321–331. https://doi.org/10.1016/j.iswcr.2020.05.002

Banister, D. (2011). Cities, mobility and climate change. J. Transport Geogr., 19(6), 1538-1546. https://doi.org/10.1016/j.jtrangeo.2011.03.009

Bergantino, A. S., Intini, M., & Tangari, L. (2021). Influencing factors for potential bike-sharing systems in urban areas: a case study from Florence, Italy. Sustainability, 13(4), 2020Apple (n.d.). COVID-19 Mobility Trends Reports, Apple. Retrieved August 28, 2020, from https://covid19.apple.com/mobility. Firstly, this paper aims to classify and cluster bike users based on their travel patterns using unsupervised learning techniques. It utilizes the Mobility Trends Reports, a dataset provided by Apple, to analyze the mobility patterns of bike users across different cities. The dataset contains data on the travel behavior of bike users, including the distance traveled, speed, and frequency of trips. The goal is to identify different clusters of bike users based on these variables, with the aim of understanding the mobility patterns and identifying potential influential factors for bike-sharing systems. The analysis is conducted using a combination of unsupervised learning algorithms, namely k-means clustering and hierarchical clustering. The k-means clustering algorithm is employed to group bike users into six clusters based on their travel characteristics. These clusters are then analyzed to identify the most significant influencing factors. The paper contributes to the growing body of research on bike-sharing systems and provides insights that can be used to improve the design and implementation of such systems.
Schelp, C. (2018). An Alternative Way to Plot the Covariance Ellipse. https://carstenschelp.github.io/2018/09/14/Plot_Confidence_Ellipse_001.html.

Seoul Institute of Technology. (2020). Technology and policy trends in the post-COVID-19 era. Seoul Technology Research, 06.

Shaheen, S. A., Mallery, M. A., & Kingsley, K. J. (2012). Personal vehicle sharing services in North America. Res. Transp. Bus. Manag., 3(June), 71–81. https://doi.org/10.1016/j.rtbm.2012.04.005

Shakibaee, S., de Jong, G. C., Alpkökin, P., & Rashidi, T. H. (2021). Impact of the COVID-19 pandemic on travel behavior in Istanbul: A panel data analysis. Sustain. Cities. Soc., 65, Article 102619. https://doi.org/10.1016/j.scs.2020.102619

Teixeira, J. F., & Lopes, M. (2020). The link between bike sharing and subway use during the COVID-19 pandemic: The case-study of New York’s citi bike. Transportation Research Interdisciplinary Perspectives, 6, Article 100166. https://doi.org/10.1016/j.trip.2020.100166

Wang, M., & Zhou, X. (2017). Bike-sharing systems and congestion: evidence from US cities. J. Trans. Geogr., 65, 147–154. https://doi.org/10.1016/j.jtrangeo.2017.10.022

Xing, Y., Wang, K., & Lu, J. J. (2020). Exploring travel patterns and trip purposes of dockless bike-sharing by analyzing massive bike-sharing data in Shanghai, China. J. Transp. Geogr., 87, Article 102787. https://doi.org/10.1016/j.jtrangeo.2020.102787

Yan, X., Yang, W., Zhang, X., Xu, Y., Bejleri, I., & Zhao, X. (2021). A spatiotemporal analysis of e-scooters’ relationships with transit and station-based bikeshare. Transp. Res. Part D: Transp. Environ., 101, Article 103088. https://doi.org/10.1016/j.trd.2021.103088

Yang, Y., Cao, M., Cheng, L., Zhai, K., Zhao, X., & De Vos, J. (2021). Exploring the relationship between the COVID-19 pandemic and changes in travel behaviour: A qualitative study. Trans. Res. Inter. Pers., 11, Article 100450. https://doi.org/10.1016/j.trip.2021.100450

Yu, S., Liu, G., & Yin, C. (2021). Understanding spatial-temporal travel demand of free-floating bike sharing connecting with metro stations. Sustain. Cities. Soc., 74, Article 103162. https://doi.org/10.1016/j.scs.2021.103162

Zhao, J., Guo, C., Zhang, R., Guo, D., & Palmer, M. (2019). Impacts of weather on cycling and walking on twin trails in Seattle. Transp. Res. Part D: Transp. Environ., 77, 573–588. https://doi.org/10.1016/j.trd.2019.09.022

Zhong, M., Hunt, J., & Lu, X. (2008). Studying differences of household weekday and weekend activities: A duration perspective. Transp. Res. Rec., 2054, 28–36. https://doi.org/10.3141/2054-04