Research on the optimization of urban residents' travel structure based on big data visualization

Wenrui Li (✉ U22092120087@cityu.mo)
City University of Macau

Ziyi Chen
City University of Macau

Xueyuan Li
City University of Macau

Su Shi
City University of Macau

Kechun JU
City University of Macau

Yiyang Li
City University of Macau

Article

Keywords: data mining, data visualization, big data analysis, public transportation, urban transportation, transportation

Posted Date: April 18th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2737117/v1

License: ☑️ This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

The travel patterns of urban inhabitants are not optimized and require improvement. By utilizing big data visualization, the optimization of urban residents' travel structure can be efficiently achieved, resulting in the reduction of urban traffic and the promotion of a healthier lifestyle for individuals. Through the utilization of big data visualization, the optimization of urban residents' travel structure can be comprehended by means of a mathematical model and the implementation of effective policies aimed at enhancing human settlement. This can also serve as a significant point of reference for exploring future studies in this field. The focus of this paper is to examine the influence of big data visualization technologies that offer insights into the routes of transportation and the movements of pedestrians, as captured by public cameras situated worldwide. In the course of transportation planning, conventional statistical methodologies and extensive data analysis techniques are employed to optimize the travel patterns of urban residents using a basic programming language such as Python. The research has confirmed the efficiency of big data visualization technologies in ameliorating the transportation framework of city dwellers.

Introduction

As the era of e-commerce proliferation continues, the urban populace's travel structure has assumed greater importance. The travel structure of urban residents pertains to the configuration of human mobility within an urban area. The subject matter under consideration concerns the various modes of transportation that people employ to travel from their residences to their places of employment, education, and other locations. Hence, the present inquiry seeks to examine and evaluate the capacity of big data visualization to enhance the mobility system of urban areas by optimizing the allocation of resources within the city.

The aim of this research is to examine the impact of big data visualization on the travel behaviors of urban inhabitants, focusing on three key aspects: (1) analyzing people's moving patterns, (2) optimizing the resources utilization on transport modes, and (3) evaluating which major problems are related to the optimized travel structure. This study will enable cities to enhance their understanding of the requirements of their residents. By utilizing big data visualization and implementing various traffic models that are tailored to the unique characteristics of each city, the study will facilitate the efficient allocation and utilization of resources. (Bliss D W, 2015). This paper, due to a lack of sufficient empirical evidence in analyzing urban residents' travel patterns and optimization of resource utilization in cities, focuses on three aspects.

1. Theoretical Background

The notion of the travel structure of city dwellers, as per Liu's (2012) perspective, is closely intertwined with transportation and its correlation with the day-to-day pursuits of people. The structure of urban residents' travel is predominantly shaped by their daily routines. The distinct routines of urban
inhabitants can have an impact on their movement behaviors. Another possible variation could be: "To provide an example, certain inhabitants of metropolitan regions may reside in closer proximity to their customary places of work or study (such as commercial centers), while others may be situated at a greater distance from these locations. Moreover, some may elect to utilize public transit while others may prefer private modes of transportation." However, in most cases, urban residents are more likely to move around the city by using public transport rather than private vehicles (e.g., cars). The travel structure of urban inhabitants is primarily influenced by the distinctive features of various subpopulations and their corresponding modes of transportation and resource employment. The population's traits in a given region have a bearing on the travel arrangement of city dwellers. For example, there is a correlation between income levels and the likelihood of utilizing public transportation, as those with higher income levels are more prone to utilizing personal vehicles for private transportation purposes rather than relying on public transportation. The transportation mode that serves those people will affect their traveling patterns. For example, the proximity of individuals' residences to railway stations may influence their inclination towards public transportation, with those residing in closer proximity more likely to utilize it. Conversely, those residing farther from these stations may opt for private vehicles instead of public transportation (Dai K, 2015).

According to Liu (2012), the concept of urban residents' travel structure is mainly influenced by the following factors: 1) the transportation model, 2) the distribution of population, and 3) the transportation mode and resources utilization model. The urban populace's travel patterns are primarily influenced by the transportation model. The populace's dispersion will determine the number of individuals who will travel in varying directions at distinct times. And each person's movement pattern will be affected by his life routine, social status, and income level. Transportation mode and resources utilization model is also key influencing factors that affect urban residents' travel structure. The key factors are the kind of transportation, distribution of population, and transport mode that serve people. This study uses the built-in parameters included in the Open Transportation Model (OTM) to analyze urban residents' travel structure. The OTM is a computer-based simulation model that replicates the transportation system, taking into account the distinctive features of transportation modes and resource allocation. Its purpose is to simulate urban transport and traffic patterns under varying conditions and scenarios. Various data sources, including location, demographics, mobility survey data, and transactional data, are employed by the OTM to compute different parameters, such as cost indicators for various modes or resources utilized by passengers, including the cost per vehicle. Therefore, in this study, the first research question is:

**In what situations will the utilization of public transportation by urban inhabitants vary?**

Urban residents' travel structure can be affected by many factors. The key factors are how long it takes to get from home to work or school, and how far they live from their workplace or school. As a result of these two factors, urban residents might take public transport rather than private vehicles to reach their destination (Drylands, 2016). Therefore, people's predefined travel mode is a key issue when it comes to their traveling behavior in urban areas. In this study, the second research question is:
In what situations will the choice of travel patterns by city residents vary?

In order to ameliorate and streamline the travel arrangement of city dwellers, it is essential to evaluate their mean travel distance and compare it with that of other urban centers. The advent of e-commerce has resulted in a considerable proportion of people living at a considerable distance from their workplace or educational institution. In light of this circumstance, the populace is devoting an increased amount of time to utilizing public transportation on a daily basis and recouping expenditures through e-commerce. Hence, it is crucial to analyze the distances traveled by citizens in order to optimize the resources of each urban center. The third research question is:

In what situations will urban residents’ travelling distances vary?

This paper’s major objective is to investigate the impact of big data visualization on urban dwellers’ travel structure, from three aspects. Firstly, the focus of this paper is on the examination of how big data can facilitate cities in comprehending the mobility behavior of their denizens and optimizing the utilization of their resources by introducing various traffic models that are customized to suit the specific features of each city. Secondly, the modal choice between public transport and private vehicles is analyzed based on different parameters, including resource allocation (space and time), accessibility cost, convenience cost, and congestion cost, in order to compare the empirical evidence. Thirdly, it analyzes the characteristics of urban residents’ traveling distance.

In what situations will the cost recovery by travel mode and travel distance vary?

The present study utilizes three traffic models to investigate the travel patterns of urban inhabitants. These models incorporate macro and micro-level variables, including city features and the lifestyles of urban residents. Finally, the results show that city residents travel between home and work or school less than expected, in terms of cost (Qin J, 2016). Therefore, in order to improve urban residents’ travel behavior and optimize their resource usage, it is worth studying the average traveling distance of urban residents as well. The final segment of this paper presents a concise summary after scrutinizing the influence of three factors on the travel behavior of urban inhabitants, specifically personal traits, cost recovery, and average travel distance.

Modal choices between lifetime income and access costs

The study utilized household-level indicators from the year 2013 as its primary data source. These indicators were derived from the classification of Chinese cities according to their respective household income levels, and the household income data was collected during the specified period. In order to collect data on urban residents’ travel behavior, household-level indicators such as occupied floor area (OA) and households with young children are used. For example, a building with a higher population density will have more people living in one house. At present, there are a total of 158 cities in China that have been categorized based on their respective household income levels [11]. To determine the household income level of each city, we have computed the average income levels of urban households.
over a period of five years. Additionally, we have obtained the average population density of these 160 cities from [12].

2. Literature Review

2.1 Overview of Urban Traffic Simulation Models in the Literature

The development and planning of a modern metropolis necessitate careful consideration of its urban transportation infrastructure. The transport modes that people choose at different times affect their travel patterns and modify their traveling time (Hughes, 2010). The phenomenon of urbanization brings about an escalation in the number of alternative modes of transportation, comprising both public transport and privately owned vehicles. Therefore, urban transportation models are essential tools to evaluate how people actually use these alternative modes (Dong et al., 2017). Many studies have suggested that the most important factors for urban residents’ travel behavior are subjected to inclement weather and financial constraints. In this case, the transport mode choice depends on accessibility cost, transport time, and convenience cost; Therefore, transport models are used to predict potential changes in travel behavior (Bagus et al., 2007). Most of these models are based on household-level indicators such as household types according to income level, population density, and self-reported commuting distance (Atsushi, 2017).

In addition to transportation systems, urban transportation models also consider other variables such as economic conditions (Hwang et al., 2006) or cultural background (Koromilas et al., 2010). However, empirical studies have shown that household-level indicators such as household income levels have an important impact on the traffic choice decision (Hwang et al., 2006; Dong et al., 2017), for example, individuals with higher household income levels use public transport more often. Dong and Li (2017) provide empirical evidence on this topic and find that household-level indicators such as household income level have a significant impact on the mode choice.

Furthermore, transportation models consider factors beyond economic constraints, including the quality of infrastructure. Therefore, urban transportation models are used to estimate the potential changes in travel behavior and estimate how well alternative transportation modes will cope with such changes (Zhang et al., 2004). In this case, transport models take into account culture (Koromilas et al., 2010) and nature (Zhang et al., 2004).

Several empirical studies have investigated and compared the modal choice based on an individual’s lifetime earnings and the costs associated with accessibility. For example, Hwang et al. (2006) first use household-level indicators to examine the mode choice decisions according to lifetime income. This paper finds that individuals with higher household income levels choose public transport more often than individuals with lower household income levels (Vpercival R, 2016).

2.2 Overview of Urban Traffic Simulation Models in the Literature
Currently, transportation models are used to evaluate different options for infrastructure investments (Deakin and Luijkx, 1993). Urban transport models are also used to estimate the potential changes in travel behavior and estimate how well alternative transportation modes will cope with such changes (Zhang et al., 2004). These urban transport models assume that land use characteristics affect the travel behavior of urban residents based on accessibility cost, transport time, and convenience cost. In this case, traffic simulation models are used in urban transport planning (Dong et al., 2016), which is used by city planners to evaluate different options for infrastructure investments (Deakin and Luijkx, 1993) under uncertainty.

The mode choice decision is influenced by many factors such as household income, public transport usage, population density, and access cost. Household income is one of the most important factors. From this perspective, Hwang et al. (2006) use household-level indicators such as household income levels to examine individuals’ transport choice decisions according to lifetime income. Individuals with higher household income levels choose public transport more often than individuals with lower household income levels.

In addition to transportation systems, urban transportation models also consider other variables such as economic conditions (Hwang et al., 2006) or cultural background (Koromilas et al., 2010). However, empirical studies have shown that household-level indicators such as household income levels have an important impact on the traffic choice decision (Hwang et al., 2006; Dong et al., 2017), for example, individuals with higher household income levels use public transport more often. Dong and Li (2017) provide empirical evidence on this topic and find that household-level indicators such as household income level have a significant impact on the mode choice.

2.3 Key Terms

Access cost: A transport cost or a cost associated with the travel time of public passenger transport systems; usually expressed as a monetary value per unit distance traveled.

Access infrastructure: A transport infrastructure that makes urban traffic flows more efficient by reducing travel time. Examples of access infrastructures include expressways and arterial roads, which provide safe and direct routes for public passenger transport and reduce travel time, as well as the development of integrated networks linking different transport modes such as bus transit, railways, metro and tramways (Feng et al., 2017); the importance attributed to these transport access systems can be seen in the measures used to evaluate them (e.g., travel times).

2.4 Big Data’s Theoretical Foundations

(1) Big data essential exploration

Top to bottom investigation of big data-related information should be completed from three viewpoints:
1) Examining primary hypothetical thoughts big data is figured out through the advancement of hypothetical thoughts. In the first place, it needs to comprehend how big data is characterized all through the business, beginning with its qualities and improvement after some time as the need might have arisen to utilize big data innovation. Also, there are additional necessities that should be met to investigate the results of big data creation, related viewpoints, and the course of movement.

2) Specialized hypothesis examination: The development of big data has ignited a specialized upset and change all by itself. Innovation, for example, distributed computing, database disseminated activity handling, capacity specialized prerequisites, and scientific models that exhibit the standards of the big data working framework fills in as the reason for big data analysis and as a device for execution.

3) Utilization of big data hypothesis in certifiable situations: The last objective is to try what you've found out about big data. This article shrewdly mixes spam SMS capture innovation and big data analysis to yield more exact vital outcomes. Big data stages will be utilized to accumulate, process, figure, and store the data during sending.

On account of the fast progression of new ideas and advances like distributed computing, data focuses, and the Web of Things as well as the continuous improvement of data innovation, the time of big data has startlingly shown up, and we are currently totally drenched in it in all features of our work, study, and regular routines. assume a vital part.

Big data has four characteristics: amount, assortment, speed, and veracity. The first is that there is a colossal measure of data. The variety of data is considered in this premise. 80% of the new data comes from sources other than the traditional data sorts. These new data should be taken care of and broken down utilizing the first model. Yet again handling and breaking down immense measures of dynamic new data rapidly is significant because of the very high-velocity prerequisites. To wrap things up, Big data is genuine, and along these lines, they can involve cutting-edge numerical methods and calculations related to the combination of different big data to deliver more helpful outcomes.

The expression "big data" alludes to a procedure for computing, handling, and dissecting data utilizing normal programming with a compelled measure of capacity and handling power. An ever-increasing number of differentiated data is going to turn into the essential data of the analysis object, requiring strong extra room, super-processing power, and an astute numerical estimation model. Handling gigantic and very enormous scope data sets with differentiated data structures, including organized, semi-organized, and unstructured data is a test. Big data analysis produces results that are more sagacious, more dependable in their capacity to advance shrewdly, and fitter for simply deciding.

While utilizing gigantic measures of data, calculations and other figuring models should be refreshed constantly. It is important to refresh the product calculation and different advancements for stream data and unstructured data types in light of the fact that the traditional data analysis mode is just proper for organized data types. This is substantially in the length of big data engineering utilized for data capacity, the board, analysis, and resulting content mining. Big data advancement should consequently stay aware
of innovation improvement and address the requirements of different application situations. In the data place, where apparently irrelevant data starts to be related and determined as distributed computing turns out to be an ever-increasing number of famous, these structures are broadened. As man-made consciousness frameworks are iteratively refreshed, big data consistently expand the incentive for individuals...

(2) Common big data technology

1) Hadoop

Hadoop comprises two fundamental components, namely HDFS and MapReduce, which together form a distributed processing framework that is widely adopted. The central tenet of this framework is the storage and retrieval of data in a distributed manner. The activity program for each distinct operation, or map, is distributed to every data hub. Following this, the activity outcomes from each hub are amalgamated into a singular reduction to produce conclusive results. The framework utilizes the broadly utilized Hadoop engineering, a direct programming establishment. Its business layer application can deal with a data set of the greater TB level lined up on a monstrous bunch comprised of thousands of machines. The design utilized a trustworthy, compelling, and versatile strategy for data analysis, incorporated various data handling capacities, and adjusted working execution.

The fundamental subproject of the Hadoop platform is the conveyed document framework HDFS, which serves as its backbone and is responsible for data access and storage. HDFS can store gigantic measures of data on which it runs MapReduce, HBase, and other sub-modules. It additionally offers quick throughput for getting to the data of the business layer. With its credits of being direct, easy to use, viable, free, and having a huge local area of allies, Hadoop has during the decade since its creation turned into the best option for some organizations to utilize distributed computing and big data.

Another significant Hadoop center task is MapReduce, a circulated data handling procedure. Engineers are able to transmit programs within a distributed group environment without the need to concern themselves with the intricacies of distribution. This is achieved by invoking and incorporating a set of generic software development libraries, which subsequently generate the desired operational outcomes. The central idea is equal registration, which makes data analysis more successful by utilizing equal handling on different machines. A work that should be done is separated utilizing maps into various more modest subtasks, which are then relegated to different hubs for comparable estimation. To acquire the eventual outcomes, use diminish to consolidate the consequences of the subtask estimations (W, 2016).

2. HBase database, second

The database chooses to use the key-value-based HBase non-relational database. It is a circulated, non-social database with segment capacity. The c/s architecture serves as the foundation for conventional relational databases. Although it has undergone several iterations of improvement, it can only handle a given amount of huge and unstructured data. It can also only actualize pooled database clusters. Based
on the Hadoop architecture, HBase is a distributed NoSQL database that is extremely scalable and available. The system is capable of managing vast quantities of unstructured data, executing parallel MapReduce operations, processing batches, and conducting specific queries.

3). Big data analysis steps

The big data analysis method consists of six basic steps:

1) Visual analysis

Data visualization is an essential prerequisite for data analysis, regardless of whether the individual is an expert in the field or an average user. By utilizing visual analysis, users can draw conclusions from the data itself, thereby making data processing and analysis more intuitive.

2) An algorithm for data mining

Visualization plays a significant role in the data mining process, where the machine acts as the core component. Data mining, data partitioning, and anomaly analysis are the primary methods used to extract information related to data mining. These data processing techniques not only take advantage of the speed advantage of processing the data but also process the data in terms of quantity.

3) Capabilities for predictive analysis

By utilizing visual analysis and data mining, prescient analysis enables data examiners to estimate and evaluate the potential progress pattern of data, while data mining provides data specialists with a complete understanding of the data.

4) Semantic hunt

The existing data is varied and disorganized, creating novel complexities for data analysis. As a result, specific data analysis techniques are essential to extract, communicate, and assess data. In view of this, a semantic motor is made that can involve reports to recover relevant data in a shrewd way.

5) Data management and quality

The greatest management techniques revolve around data management and quality. By following the established methodology and utilizing expert data processing tools, the acquired findings will maintain a level of quality that aligns with the predetermined targets.

6) Data warehouse and storage

To enable sophisticated scrutiny and multi-dimensional exhibition of data organized in a particular format, a social database in the form of a data distribution center has been created. The way to accomplish the intellectualization of business frameworks is through the structure of data distribution centers. The fundamental player who incorporates the data from business frameworks and supports
business insight with data extraction, soil change, and stacking (ETL). The framework's standard activity provides the groundwork for a plethora of data-related queries and framework accessibility, thereby creating a pivotal stage for collaborative multi-computational data analysis and mining.

3. Methodology

This section includes descriptions of the study location, the data-gathering procedure, and the analytics used. All methods were carried out in accordance with relevant guidelines and regulations that adhered to the Declaration of Helsinki. All experimental protocols were approved by City University of Macau. By means of a survey and social media analysis, the authors explored the way in which tourists in China interact with the inhabitants of urban areas. Which metropolitan center has become a sought-after destination? Four universities surround the transit facility, making it a popular destination for China's youth. The selection of a survey as a customary data-gathering method and social media as a big data source was a judicious decision. In contrast to other established methods, surveys enable the collection of larger samples and yield highly specialized data that is meticulously produced by researchers (Chen C, 2016). The exchange of ideas and daily activities among social media users is facilitated by social media platforms, which also provide a considerable sample size for data analysis. Researchers have conducted several studies comparing data from surveys and social media. For comparison, this study chose survey and social media data. An on-site survey was carried out at the urban dwellers on weekdays and weekends. During this time, tweets mentioning the name of the transport structure were also gathered. The results were derived from surveys and social media and which are analyzed by using statistical analysis and sentiment data.

3.1 Study Site

After taking into account the aforementioned urban population, the research location was determined. Over the course of the last century, this metropolitan resident has functioned as a railway (refer to Fig. 1). Following the conversion of the railway into an underground system, the location was left deserted and abandoned. The Seoul Metropolitan Government took up the initiative to construct the travel infrastructure on the vacant site as a part of its urban regeneration program (Kaur et al., 2020).

As a result of its present location, the populace dwelling in urban regions has garnered a significant following among the younger demographic. The district has established itself as a desirable urban hub for the younger generation, boasting a plethora of popular hangout spots.

3.2 Data Collection

The data was collected through two distinct methods, namely a survey and social media, as depicted in Table 1. The survey data was obtained through an onsite survey, with a total of 180 samples being gathered. This paper identifies all methods as ethically sound according to the statements made in accordance with relevant guidelines and regulations. 1023 tweets were filtered after collecting posts that utilized keywords related to urban residents on social media, as reported by (Xiong Y,2018).
3.3 Survey

Following convention, a survey was executed to contrast social media data, appraise the behavior of urban inhabitants who utilize these platforms, and determine their level of satisfaction with their residential location. Informed consent was obtained from all subjects and/or their legal guardians for all investigation. Surveys are one of the more conventional techniques that can satisfy both efficiency and fairness (Xiaofeng M, 2013). The ideals of democracy, representativeness, openness and public acceptance are all contingent upon the notion of impartiality. This idea focuses on how the public and participants see things and whether or not the public participation process effectively reflects the target population's viewpoints. Efficiency is a term used to describe how simple it is to collect data.

From September 18 to 22, and from October 1 to 8, 2019, the survey was carried out. Weekdays and weekends were included in the Urban Residents' sampling dates (Uzoka F M E, 2017). The researcher approached every third adult visitor to the urban area and inquired if they would be interested in participating in the survey. Out of the total number of respondents, 180 individuals confirmed their visit and willingness to partake in the poll of urban inhabitants.

The questionnaire was categorized into five subsections: respondent information, visit satisfaction, social interactions, social cohesions, and frequency of visits by urban inhabitants. Information on visits to urban inhabitants (frequency, duration, and company) and the things visitors did there was included in the first section (Chourabi H, 2012). Nine urban inhabitants' activities were thoughtfully chosen based on prior literature. User satisfaction included inquiries on the visitors' perceptions of the amenities offered to urban dwellers, with responses ranging from "1 = strongly disagree" to "5 = strongly agree." Participants were asked to rate their level of safety, participation in social programs for urban inhabitants, and ability to build new relationships when it came to social engagement. The social cohesiveness portion of the survey questioned the nature of trustworthy ties with people in their neighborhood since social cohesion depends on trusting social relationships, such as a sense of community. In line with Perez et al. (2015), survey respondents were asked to self-report on local residents' social cohesiveness characteristics and their desire to take action for the betterment of the community (Energy, 2015). The authors also added a question regarding the inhabitants' attitudes toward and readiness to form social connections with other visitors of the urban residents based on questions from the literature.

4. Results
The findings were divided into three parts, namely social media, the survey, and a contrast between these two techniques. The survey findings suggest that tourists occasionally sought repose in the presence of city dwellers. Social engagement and user happiness were positively correlated. According to the survey’s findings, urban dwellers serve as a hub for social contact and cohesion and offer a place to unwind (Kim GH, 2014). The social media data showed that urban people have favorable sentiments about other urban residents, especially when it comes to characteristics of urban residents like “forest” and “railroads (Yin J, 2013).” The findings obtained from the social media analysis indicate that guests of urban residents were inclined towards group interactions and responded favorably. In conducting a comparative analysis between survey responses and positive feedback on social media, it was found that there were no inconsistencies in terms of the levels of satisfaction expressed by the respondents.

4.1 Survey Results

The results of the poll showed the characteristics of urban residents who travel. Similar to other studies, urban residents tended to receive more visits from women (54.24%) and those under the age of 40 (64.41%). The bulk of tourists came from other districts, and about 30% of them were people who lived in nearby communities to the urban residents (Table 2).

<table>
<thead>
<tr>
<th>Demographics</th>
<th>N</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>72</td>
<td>27.4</td>
</tr>
<tr>
<td>Male</td>
<td>83</td>
<td>53.6</td>
</tr>
<tr>
<td>Female</td>
<td>38</td>
<td>38.4</td>
</tr>
<tr>
<td>Age 18–20</td>
<td>73</td>
<td>37.57</td>
</tr>
<tr>
<td>Age 25–30</td>
<td>73</td>
<td>37.57</td>
</tr>
<tr>
<td>Age 35–40</td>
<td>28</td>
<td>53.5</td>
</tr>
<tr>
<td>Age 45–50</td>
<td>127</td>
<td>39.4</td>
</tr>
<tr>
<td>Residents</td>
<td>Yes</td>
<td>29.3</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>39.4</td>
</tr>
</tbody>
</table>

The survey’s descriptive statistics outlined the fundamental purposes of urban dwellers. An explanation of the residents is provided in Table 3. 39.0% of visitors to urban residents less than once each month, and 22.8% of visitors were first-timers (Mohamed N, 2014). Additionally, a certain percentage of travelers paid weekly residential of more than two days (15.2%) or daily visits (10.8%) to urban residents. This
finding confirms that urban residents only occasionally use the urban residents and that travelers to the urban residents typically come for special events. Figure 2 shows the characteristics of travelers. Figure 3 shows the frequency of residents.

### Table 3

<table>
<thead>
<tr>
<th>Number of urban residents</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Everyday</td>
<td>10.8%</td>
</tr>
<tr>
<td>More than 3 day a week</td>
<td>15.2%</td>
</tr>
<tr>
<td>Once a week</td>
<td>39.0%</td>
</tr>
<tr>
<td>Less than once a month</td>
<td>22.8%</td>
</tr>
</tbody>
</table>

### 4.2 Comparing the Survey and Big Data

The data obtained from both the survey and social media platforms demonstrate that urban residents typically have a pleasant experience. Despite significant variations between the respondents surveyed and the urban residents who utilize the space, the general consensus was that the experiences there were favorable. This is substantiated by the factor correlation analysis presented in Fig. 4.

By analyzing both social media data and survey responses, it has become possible to gain a deeper understanding of the actions, emotions, and social interactions of urban residents. These two sources of information have provided a nuanced and comprehensive view of user behavior (Khan M, 2014). For instance, users' remarkable daily activities like sharing great meals with someone could be captured in social media data. The examined data encompassed the habitual and standard conduct of metropolitan residents, including promenading, decompressing, and conversing. Both approaches substantiate the proposition that urban denizens could enhance social engagement by means of social interaction. Social media statistics and survey results imply that urban dwellers are a location to strengthen social relationships in terms of social interactions. According to the survey, it was reported users of urban areas emphasized the role of urban residents in fostering social interaction and social cohesion. The study also revealed a positive correlation between social contact and satisfaction. Nevertheless, the factors that give rise to the correlation between the two variables have yet to be determined.

### 5. Conclusions

The objective of this investigation was to fathom the incentives underpinning the pertinence of the utilization of surveys and social media by urban residents. Big data and urban residents are two contemporary and crucial ideas; Numerous individuals have commenced merging these elements to produce intelligent urban applications that will facilitate the attainment of sustainability, enhanced adaptability, streamlined administration, upgraded living standards, and astute regulation of transportation infrastructure resources (Su K, 2011). In our research, we conducted an analysis of
concepts and their respective definitions, ultimately identifying certain commonalities between them. Notwithstanding the existence of numerous definitions, each notion exhibits a particular set of characteristics that set it apart. These shared traits enabled us to deduce the overall advantages of utilizing big data in the creation and assistance of journey-planning applications for urban dwellers. As per the poll's results, the predominant use of urban locales by visitors was for the purpose of decompressing and replenishing their energy levels from their everyday existence. Urban inhabitants who reported satisfaction with their living conditions also experienced heightened social cohesion and enhanced social interaction. According to the poll's findings, individuals residing in urban areas serve as a hub for social engagement, communal bonding, and a location for relaxation. Social media data suggests that those living in urban settings possess positive attitudes towards their peers who also reside in urban areas, particularly when it comes to qualities such as “forest” and "railroads." Based on the social media analysis, it was observed that group interactions were positively received by city visitors. Upon conducting a comparative analysis of the two sets of data, it was determined that there were no discernible disparities between the satisfaction ratings provided by survey participants and the positive comments shared on social media platforms (Lee CH, 2013).

**Declarations**

**Data Availability**

Data sets generated and/or analyzed during the current study are not publicly available due to data for academic use only and not for commercial use, but are available from the corresponding author upon reasonable request.

**Declaration of competing interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**References**

Figure 1

Description of Data Collection
Figure 2

Characteristics of Travelers
Figure 3

Frequency of residents
Figure 4

Factor correlation analysis. sati, short for satisfaction; soc, short for social interaction; and soco, short for social coherence.