

Mobile data for studying public space, and trips in Mexico City: a study case of six suburban zones

Alejandro Saniger¹, David López² , Constanza Delón¹ and Oscar Ruiz³

¹Independent Researcher, Mexico City, Mexico

²Instituto de Ingeniería, Universidad Nacional Autónoma de México, Mexico City, Mexico

³Dat's Why, Mexico City, Mexico

Corresponding author: David López; Email: dlopezfl@iingen.unam.mx

Received: 29 May 2023; **Revised:** 05 December 2023; **Accepted:** 10 January 2024

Keywords: mobile data; public space; trip patterns

Abstract


This study uses anonymized GPS traces to explore travel patterns within six suburban zones and a central area in Mexico City. The descriptive analysis presented in this paper profiles trips by distance and investigates their distribution within each zone. It examines the prevalence of local trips, walkability, and the availability and spread of entertainment sites within 15-min isochrones accessible by foot, bicycle, transit, and private vehicle. Notably, the central zone boasts diverse entertainment offerings, commendable walkability, and a substantial proportion of short and long trips. It is found that GPS traces are within their home. However, the share of long trips for the inhabitants of central zones is considerably more significant than that for the suburbs. The study highlights suburban zones that could benefit from governmental intervention to enhance transportation and pedestrian conditions. Additionally, it identifies other suburban zones that resemble the central areas in terms of walkability, trip distribution by distances, and the accessibility of entertainment places.

Policy Significance Statement

The relevance of this research can be applied as an instrument to prioritize places for interventions according to population needs. Different tools have been developed and can be customized to analyze cell phone data. This analysis is valuable for stakeholders and government decision-makers to support their strategies. Suppose public administrations seek to integrate, propose service systems at a metropolitan level or change land uses. In that case, they can base their programming on the results obtained in this study. Also, this methodology could be adapted or replicated for other types of relevant projects where planning is based on verifiable data, which will reduce execution costs and guarantee a substantial improvement in the mobility of the inhabitants.

1. Introduction

The data gathered by companies often involves tracking individuals' locations, which are used for various purposes. However, this information is anonymized to uphold privacy standards and adhere to local

 This research article was awarded Open Data badge for transparent practices. See the Data Availability Statement for details.

© The Author(s), 2024. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial licence (<http://creativecommons.org/licenses/by-nc/4.0>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original article is properly cited. The written permission of Cambridge University Press must be obtained prior to any commercial use.



regulations like the General Data Protection Regulation (GDPR). As a result, this anonymized data lack personal identifiers such as gender, age, or specific mobile brand information. Moreover, the geographical precision is intentionally reduced to block or zone centroids to safeguard privacy further. Despite these privacy-oriented adjustments, this anonymized dataset is valuable for extracting insights into people's mobility patterns, particularly when integrated with complementary datasets. Governments and decision-makers can leverage this anonymized data to support local policy initiatives and identify prevalent urban challenges.

In this study, we investigate the travel patterns, focusing on travel distances and local trips across six suburban zones and a central zone in Mexico City. By employing anonymized GPS traces, our analysis aims to discern suburban residents' preferences between long-distance travel and local trips within their immediate vicinity (local trips).

To comprehend local trip dynamics, we consider the availability of nearby entertainment spots and the overall walkability in each zone. The presence of entertainment venues could exert a notable influence on local trips compared to other travel motives. Notably, work and educational zones are more concentrated in the central region of Mexico City, leading suburban residents to often undertake longer trips for these purposes. The city boasts extensive coverage of convenience stores catering to basic necessities such as food and household supplies so that people can travel far or close, and it is most likely to have a store close to them. However, the diversity and spread of entertainment venues across different zones within Mexico City vary considerably. Hence, the provision of entertainment destinations might impact local trip behaviors.

To know the accessibility of recreational sites within each zone, we draw on the concept of the 15-minute city (Jacobs, 1992) and employ the walkability indices established by Gutiérrez-López et al. (2019), Leticia et al. (2022), and Institute for Transportation and Development Policy (2018). By delineating 15-min isochrones by foot, bicycle, transit, and private vehicle around the zones, we gauge the diversity and distribution of recreational venues within these isochrones. The six suburban and central zones were chosen using purposive sampling based on the team's familiarity with these areas in terms of transit accessibility, walking conditions, and the availability of entertainment services.

The rest of the paper is organized as follows: related works on using mobile phone data to describe the relational dynamics of individuals are presented in [Section 2](#). [Section 3](#) presents the data description acquisition for this study. We use open access data to characterize the study area, and the GPS traces from mobile phones are used to study the trips originating at the studied zones. The study area is presented in [Section 4](#). This section computed the diversity and distribution of the recreational places in each studied zone. Two indexes are constructed to know the diversity and distribution of recreational places within each isochrone for each study area. Using the GPS traces collected from mobile phones in [Section 5](#) presented the dynamics of the trips originated in each zone. In this section, two indicators are developed to know how far or close the inhabitants of each zone travel. The limitations of the study are presented in [Section 6](#). Finally, [Section 7](#) presents the conclusions and future work.

2. Related work

Pentland (2015) offers definitions for engagement and exploration, which have served as guiding principles in this paper. Engagement pertains to establishing behavioral norms and the societal pressure to uphold them, while exploration fosters innovation and social networks.

While gauging a city's level of exploration can be partially assessed by examining the frequency of long or short trips, determining the level of engagement in a specific zone poses a significant challenge. We posit that entertainment venues foster social interactions (engagement); individuals tend to engage more openly with their neighbors in such venues, compared to encounters during work hours, where stress levels might inhibit open conversation. Although this research does not explicitly evaluate the six suburban zones in Mexico City regarding exploration and engagement as delineated by Pentland (2015), both concepts have inspired the analysis presented in this paper.

GPS traces obtained from mobile phones have been widely used to describe mobility patterns and social interactions. Researchers have extensively utilized this data to gauge green space accessibility and usage (Guo et al., 2019; Xiao et al., 2019; Heikinheimo et al., 2020), study travel behavior (Jiang et al., 2017; Wang et al., 2018), and explore tourism trends (Zhao et al., 2018; Reif and Schmücker, 2020).

Studies have also examined social interactions (engagement) using geographical locations obtained with mobile phones. Eagle et al. (2008), and their follow-up work (Eagle et al., 2009) explored the potential of mobile phone data to study the proximity among individuals. By using self-reported data about friendships, the authors found that it is possible to determine whether two individuals are friends simply by analyzing their mobile phone locations. In another study, Eagle et al. (2009) were able to predict job satisfaction using the same dataset. Shi et al. (2015) employed mobile data locations to detect inner communities in Harbin, China. Ebrahimpour et al. (2020) used GPS location data reported by the social network Weibo to study human mobility patterns for smart-city planning decisions. Siła-Nowicka et al. (2015) collected GPS traces from 250 participants to examine how travel mode depends on residential location, age, or gender. Additionally, they identified the so-called *third places* beyond the primary locations (home and work) where individuals socialize. Finally, Takaki (2018) used anonymized GPS data from mobile phones and social media posts to demonstrate how inhabitants of Sao Paulo, Brazil, naturally appropriate physical and virtual spaces.

The literature presents several recent studies examining travel patterns and their implications for stakeholders and policymakers; some recent works are presented as follows. Katrakazas et al. (2020) examined COVID-19's impact on driving behavior and safety metrics using a dedicated smartphone app. Their findings suggest policymakers should focus on implementing revised speed limits and creating more extensive spaces for cyclists and pedestrians. These measures aim to increase distances between road users, thereby promoting heightened road safety and minimizing the spread of COVID-19. Chand et al. (2021) also studied COVID-19 impact on driving behavior; in their study, the authors analyzed the traffic incidents duration and frequency during the pandemic and found a shift in the location of incidents, with more incidents recorded in suburban areas, away from the central business area. Zou et al. (2020) presents statistical data on the e-scooter trip distribution and then conducts an exploratory examination of trip paths alongside street-level characteristics. The authors delve into policy and regulatory considerations concerning e-scooter infrastructure design and safety concerns. Finally, Badawi and Farag (2021) analyses Saudi women's travel patterns from 2015 to 2020. It highlights that young Saudi women predominantly rely on private cars for mobility. Despite their interest in adopting sustainable transportation, there has been no statistical shift in travel behavior over the past 5 years. The study suggests that the government should establish a women-friendly urban environment to promote sustainable mobility among Saudi women. This environment should encompass a varied, well-connected, affordable, safe, and flexible mobility network that links crucial destinations and amenities.

The literature discussed in this section highlights GPS traces as a potent tool for extracting diverse travel patterns within a city's populace. Additionally, conducting descriptive data analysis enables the identification of various issues, providing valuable insights for decision-makers seeking to implement effective policies.

3. Data

In this study, two critical open-access datasets are utilized. Firstly, the national directory of economic units serves as a comprehensive repository containing all economic entities nationwide. Notably, it records entertainment establishments operating within the country, categorizing them as public or private venues. The directory encompasses various entertainment enterprises, ranging from bowling alleys and water parks to museums and theatres (Instituto Nacional de Estadística, Geografía e Informática, 2023).

Conversely, the national census conducted in 2020 spanned from March 2 to March 27, engaging over 147 thousand interviewers. Covering nearly 2 million square kilometers of the national territory, this extensive effort involved visiting each housing unit to gather comprehensive demographic, socioeconomic, and cultural data about the Mexican populace (Instituto Nacional de Estadística, Geografía e

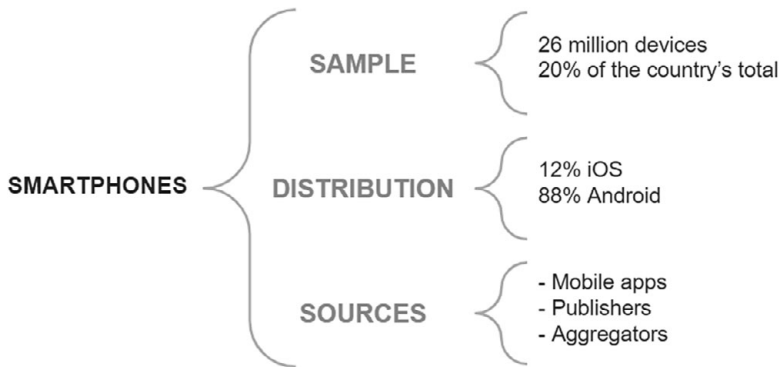


Figure 1. Distribution of devices where *Dat's Why* can collect information.

Informática, 2020). These open-access datasets formed the foundation for comparison and contrast with the mobile phone-derived data in this study.

3.1. *Dat's Why* mobile phone data

Dat's Why is a company committed to understanding the geographical behavior of individuals within urban areas through the utilization of big data and analysis algorithms. Through an extensive network of millions of interconnected devices, the company gathers data on the mobility patterns of individuals, vehicles, roadways, and transportation systems across Mexico and some other Latin American regions.

In this research, *Dat's Why* GPS traces, acquired from mobile phones, have been employed to analyze travel patterns from people in the study areas. These GPS traces¹ are obtained through the integration of a software development kit (SDK) within mobile applications. When a user interacts with an app that includes the SDK, the phone periodically transmits the user's location data.

Due to privacy restrictions, it is impossible to determine which application is running, the application category (e.g., gaming, finance or social network) or any additional private information related to the user. The only information available is the GPS location, time of the day, and the operating system (OS; Android and iOS). However, it is impossible to know the OS version due to privacy restrictions. Figure 1 shows the distribution of devices where *Dat's Why* can collect information; the share of OS of the collected data is 88% Android and 12% iOS. According to StatCounter (2023), as of 2019, Android market share in Mexico was 85.6%, while iOS was 14.0%. This share is similar to the data collected by *Dat's Why* from Mexicans.

Dat's Why infers information about the home location for the individuals from whom they collect data at a broad level; due to privacy protection laws, it is not done in more precision. The process of inferring the home area is as follows.

Before delving into the home zone inference, it is important to understand how geographic areas are divided. The zones are determined based on the Geostatistical Basic Area (AGEB) classifications used by the National Bureau of Statistics and Geography. An urban AGEB represents a defined geographical area composed of blocks precisely bordered by streets, avenues, walkways, or other easily identifiable ground features. These areas are predominantly characterized by residential, industrial, service, and commercial land uses. Urban AGEBs are assigned within urban areas, defined as areas with a population exceeding 2,500 inhabitants and municipal capitals. A typical urban AGEB spans over 25 to 50 blocks.

Dat's Why infers the home location at an AGEB level by identifying the most probable living AGEB for each monitored device. Identifying the home zone involves determining where a device exhibits

¹ These GPS traces, hereafter referred to as traces, represent location data generated by a device at specific moments through the SDK (software development kit).

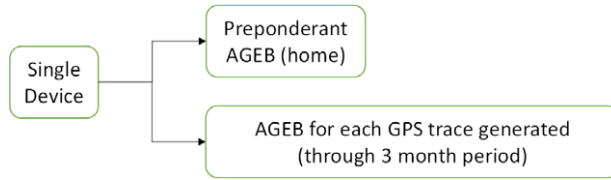


Figure 2. *Dat’s Why* data management.

Table 1. Database dictionary

Field	Description
device_id	Unique ID assigned to specific mobile phone
stat_area_2	AGEB where the GPS trace is generated from
total	Total number of GPS traces from same device ID, generated at specific AGEB
preponderant	Home AGEB assigned that specific device
distance	Distance from device’s preponderant AGEB to trace’s AGEB
distance_class	Distance category: <i>Proximity</i> (<2 km), <i>intermediate</i> (<5 km), <i>distant</i> (>5 km)

activity between 1:00 a.m. and 6:00 a.m. The procedure entails identifying the AGEB with a high density of data points during this time frame. Once an AGEB is recognized as a potential candidate for a device’s home, it undergoes validation to confirm whether the point aggregation repeats the same AGEB and period for at least 2 days in a given month. These two days do not need to be consecutive. The 2-day criterion is based on the lifespan of each device’s unique identifier and the frequency of location reporting. This 2-day period minimizes the risk of erroneously assigning homes while retaining many devices in the analysis. On average, a device’s unique identifier has a reporting frequency of 2 to 6 positions throughout the day and 1 to 4 at night.

3.2. Database structure

Dat’s Why collects data from individual devices and assigns each device to a predominant AGEB (home AGEB). All GPS traces collected from the device during 3 months (January–March 2019) are then correlated to a device, as shown in Figure 2. Note that a GPS trace could traverse several AGEBs throughout the day, indicating movement or visits across different geographical areas within the 3 months.

The collected data from *Dat’s Why* was used to create two databases for this study. The initial database comprised one row for each GPS trace. Subsequently, an aggregated database was generated by consolidating all records from the same device within the same AGEB. The analysis for this study was conducted using the aggregated database, and detailed field descriptions can be found in Table 1.

4. Study area

In considering the socioeconomic and spatial attributes of the metropolitan area of Mexico City as a reference, our analysis focuses on seven distinct zones, comprising six suburban areas and a central zone. Our study zones were delineated based on three Geostatistical Basic Areas (AGEB) to comprehensively represent the diverse dynamics of interaction, mobility patterns, and transfer times. These zones were explicitly chosen to encompass areas with an average population of approximately 20,000 inhabitants while also exhibiting varying socioeconomic characteristics. Additionally, their geographic locations were selected to contrast the central area and the suburban zones. Among the six suburban zones under consideration, three are located within Mexico City, while the remaining three fall within the State of

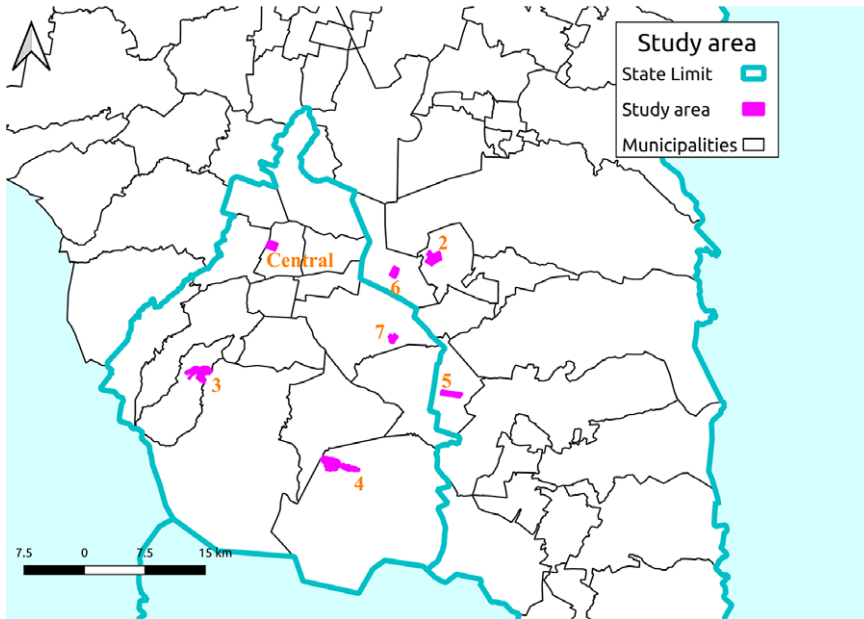


Figure 3. Study area.

Table 2. Population and the number of unique devices that collect GPS traces

Zone	Name	Population	Total devices
Central	San Rafael	22,567	29,684
2	Chimalhuacán	24,542	4,471
3	M. Contreras	26,829	2,792
4	Milpa Alta	21,015	1,383
5	Chalco	18,422	1,805
6	Neza	21,099	4,561
7	Iztapalapa	18,043	1,899

Mexico. It is worth noting that accessibility to the central area from the suburban zones in the State of Mexico might pose a more significant challenge due to a comparatively lower level of integration within the transportation systems.

Figure 3 provides a visual representation of these defined zones, each with its unique attributes as detailed below:

- Central zone: is located in the downtown area of Mexico City.
- Zones 3, 4, and 7: are the suburban zones situated in Mexico City.
- Zones 2, 5, and 6: are the suburban zones located in the State of Mexico.

Dat's Why database was utilized directly to obtain 3 months' worth of data for all zones. Complete available datasets were used for each zone, and the distribution of OSs for each zone is, on average, 12% iOS and 88% Android. In Table 2, the total population of each zone and the count of unique devices with identified home locations in each zone are provided. All devices collected by *Dat's Why* with the corresponding preponderant AGEb (home AGEb) for each zone were utilized without employing any

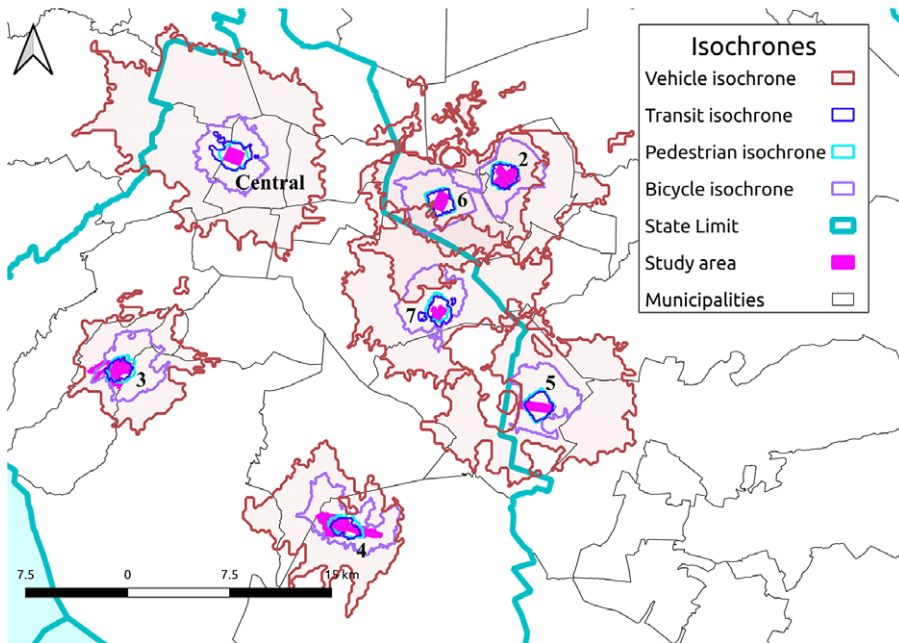


Figure 4. 15-min isochrones using vehicle, transit, bicycle, and walking.

sampling method. The central zone is a control area with excellent accessibility to various transportation modes. Zones 2–7 are serviced by a maximum of two transportation modes, lacking mass transit (subway and BRT) within a 15-min walk.

Table 2 demonstrates that in suburban zones, the average portion of devices from which information is collected represents 15.6% of the total population, whereas in the central zone, this share extends to 130%. It is pertinent to note that individuals with higher incomes, typically found in the central zone, may possess multiple or newer smartphones capable of accommodating more applications, potentially increasing the likelihood of having the specific SDKs utilized by *Dat's Why* for GPS trace collection. However, as discussed in Section 5, despite variations in database sizes across zones, patterns such as the proportion of individuals travelling long distances or remaining at home are comparable between zones, attributable to the substantial dataset sizes in all cases.

In each zone, 15-minute isochrones are computed for travel by foot, bicycle, transit, and private vehicles. The Valhalla routing service is used to compute the pedestrian and bicycle isochrones (Valhalla, 2023), while TravelTime API (TravelTime, 2023) was used to compute the transit and vehicle isochrones. These routing services leverage the information from Open Street Maps for computing paths in any city. In order to not overestimate or underestimate the travelling times for each mode, real-life parameters such as walking/biking speed, hills, and pavement conditions, among others, are taken into account. The Supplementary Material describes all the parameters used for the isochrone computation.

Figure 4 shows the isochrones for each mode. Figure 5 shows the isochrone areas for pedestrian, bicycle, and transit isochrones. The vehicle isochrone area is not shown in Figure 5 for correct figure visualization. In the Supplementary Material, the areas of the isochrones for all modes are found. By foot, a person can cover the same area in all zones. However, the coverage area varies significantly for the rest of the modes. The area that can cover the inhabitants of Zones 4 and 5 by bicycle or vehicle is significantly smaller than in the rest of the suburban zones. In these zones, the network connectivity is bad compared to other zones.

Transit in the metropolitan area of Mexico City lacks schedules or official frequencies, so it is hard to compute accurate waiting times for transit isochrones. The transit administered by the Mexico City

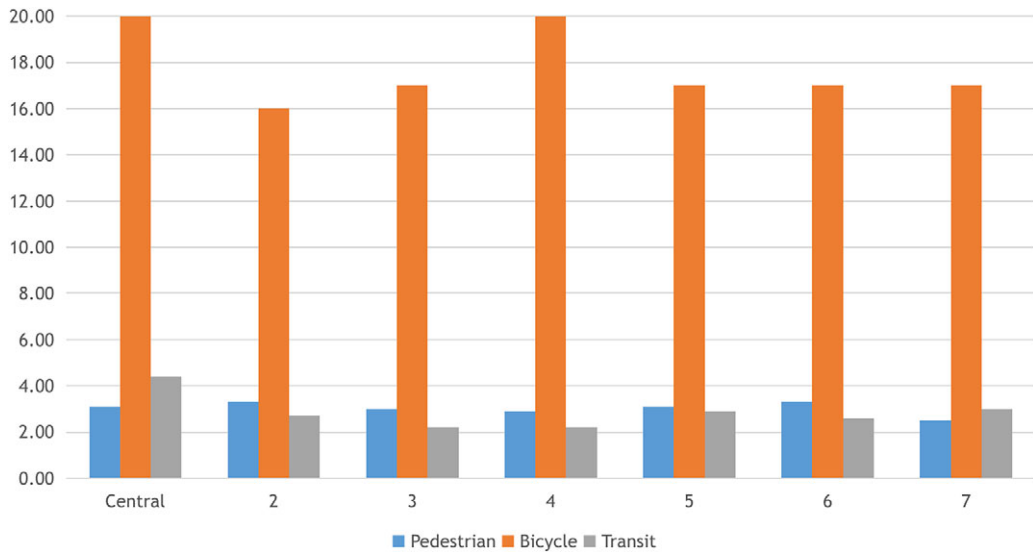


Figure 5. Isochrone areas for the pedestrian, bicycle, and transit. For visualization purposes, the area of the automobile isochrone is not shown.

Government has GTFS information available. However, it is difficult to trust the frequencies reported in the GTFS. The only reliable information in the GTFS is the routes and stops. Also, around 80% of the transit routes are small franchises with little government supervision. So, there is no reliable information about these franchises' routes, stops, and waiting times. In particular, no public information exists on transit routes for Zones 2, 5, and 6. The only well-grounded information for transit is found in the central zone and Zones 3, 4, and 7. Although there is no information for franchise transit in these zones, we have data for the transit operated by the government. It is noted that in Zones 3, 4, and 7, the pedestrian and the transit isochrones are almost equal in area, so the reach by transit (operated by the government) of the inhabitants of these zones does not reach further away in comparison to the pedestrian trips.

The Herfindahl–Hirschman (HHI) and Entropy (EI) indices, drawn from Song et al. (2013), are employed to evaluate the diversity and distribution of recreational places within each isochrone across different zones. The HHI, ranging from 0 to 10,000, gauges the diversity of recreational places within each isochrone. Lower values signify greater diversity, while higher values suggest a lesser mix. For instance, an isochrone featuring a blend of gyms, theatres, and parks would have a lower HHI than one with solely gyms and theatres.

The EI measures the distribution of recreational places within an isochrone. Ranging from 0 to 1, an EI of 1 indicates an ideal, balanced distribution, whereas an EI of 0 reflects an unbalanced arrangement. For instance, an isochrone dominated by 80% gyms and 20% theatres would yield a lower EI than one with an equal split of 50% gyms and 50% theatres. For further details, refer to the Supplementary Material for the mathematical formulation of HHI and EI and the list of considered recreational places.

Upon assessing the diverse zones, it is evident that the central zone exhibits the highest HHI for pedestrian, transit, and private vehicle isochrones, signifying the most varied recreational places among zones. However, the central zone shows a less balanced distribution for vehicle and bicycle isochrones, reflecting a lower EI. Notably, the majority (85%) of recreational places in these isochrones consist of gyms, museums, and cultural venues.

Zone 2 excels in diverse recreational places among suburban zones, notably outperforming the central zone in bicycle isochrones. Additionally, it boasts the most balanced distribution of recreational spots for pedestrian and transit isochrones across all zones.

Zones 4 and 5 exhibit a good distribution of recreational spots for bicycle isochrones but lack a diverse mix, that is, in these zones for the bicycle isochrone, there are few options for recreation, but the options are well distributed. Zones 3–5 and 7 in pedestrian and transit isochrones feature limited or no entertainment places within 15 min, necessitating inhabitants to travel farther for recreation. The authors are aware that the previous assumption might not be valid for transit since, in this research, the transit network is incomplete, given the lack of public and reliable information.

Regarding private entertainment, the central zone exhibits the greatest diversity across all isochrones but lacks adequate distribution. Approximately, 75% of recreational venues in the central zone, across all isochrones, are primarily arcades, gyms, and venues hosting artistic, cultural, and sports events. Zone 3 shows potential for better distribution and variety in public entertainment centers for pedestrians and transit isochrones. Observe that Zone 3 excels in private entertainment distribution. Assuming that demand shapes the distribution of private entertainment, Zone 3 could benefit from increased public entertainment venues. This observation hints at an opportunity for government intervention to introduce and improve public entertainment facilities. Although this observation could identify a candidate for government intervention, more studies would be required if the government wanted to assign a budget for such improvement. Zones 4–7 perform poorly regarding diversity and distribution of private entertainment venues, where nearly 90% of these places are either gyms or arcades. [Table 3](#) summarizes the diversity and distribution of recreational places for each zone.

5. Trip analysis

The mobile phone dataset comprises 46,000 anonymous devices, and over 3 months, all devices reported 1,000,000 traces. Each device is assigned to a unique home-AGEB in the study area using the method described in [Section 3.1](#), and it is assumed that each device belongs to a unique individual during the 3 months. [Table 4](#) shows the number of unique devices in each zone and the total GPS traces reported by all those devices. For instance, in Zone 2, 4,471 devices reside in this zone, which generates 111,834 traces in the metropolitan area.

This section presents a descriptive analysis of trips within the mobile phone dataset. To determine the distance from home a device is during the 3 months, the AGEBS in the metropolitan area where a device reports a location is selected. Then, an indicator is developed to determine how far or close a device is to its home-AGEB. This indicator is used to assess whether a given zone's inhabitants travel close to their homes or further away. Subsequently, the devices are profiled to ascertain whether they undertake more local trips (less than 2 km) or longer-distance trips (more than 2 km). At the end of the section, walkability indicators are obtained to examine the correlation between the number of local trips and the walkability of each zone.

As discussed in the limitations section of this paper, it could be possible to complement this information with the OD survey for the metropolitan region of Mexico City, which has disaggregated data for gender and different age groups. It was not done because the study areas in this paper are considerably smaller than the OD districts, and the behavior at the district level would not be representative of smaller specific areas.

5.1. General heat maps from mobile phone data

An initial origin-destination analysis is made for each zone. Given the home-AGEB of each device and GPS traces, all the AGEBS a device visited in the 3 months from 6:00 to 22:00 hrs are obtained. [Figure 6](#) shows a heat map with all the AGEBS the devices touch throughout the day. In [Figure 6](#), the black AGEBS indicate that 75% to 100% of the devices originating in a given zone were found at least once in such AGEB for the 3 months from 6:00 to 22:00 hrs. The purple AGEBS indicates that between 75% and 100% of the devices with origin in a given zone visit the AGEB at least once. The orange AGEB indicates the destination between 25% and 50% of the devices originating in a given zone. The light yellow AGEBS are for destinations between 0% and 25%.

Table 3. *The green color indicates the best performance for each index, while red indicates the worst*

Isochrone	Study zone	Entropy index			Herfindahl–Hirschman index		
		Public entertainment	Private entertainment	Public and private	Public entertainment	Private entertainment	Public and private
Pedestrian	Central	0.25	0.21	0.17	3,057.85	2,583.83	1,720.00
	2	0.21	0.21	0.18	5,555.56	4,178.15	3,549.38
	3	0.00	0.30	0.29	10,000.00	3,472.22	2,622.22
	4	0.00	0.29	0.29	10,000.00	5,000.00	3,877.55
	5		0.27	0.27		4,733.33	4,733.33
	6	0.00	0.25	0.23	10,000.00	5,292.97	4,723.18
	7		0.25	0.25		4,844.44	4,844.44
Bicycle	Central	0.26	0.16	0.13	3,304.50	2,443.41	1,829.80
	2	0.29	0.12	0.08	1,875.00	4,632.69	4,124.49
	3	0.25	0.18	0.15	4,844.44	4,898.87	3,616.75
	4	0.34	0.23	0.23	5,061.73	3,786.85	2,311.11
	5	0.35	0.19	0.15	5,000.00	4,868.21	4,610.58
	6	0.23	0.11	0.09	3,079.58	4,935.26	4,120.23
	7	0.30	0.17	0.12	3,333.33	4,413.96	3,970.79
Transit	Central	0.24	0.20	0.17	3,367.35	2,241.74	1,486.33
	2	0.21	0.24	0.21	5,555.56	4,362.24	3,610.82
	3	0.00	0.32	0.32	10,000.00	4,400.00	3,061.22
	4		0.25	0.25		6,800.00	6,800.00
	5		0.27	0.27		4,759.95	4,759.95
	6	0.00	0.25	0.24	10,000.00	5,336.08	4,673.01
	7		0.24	0.24		5,069.25	5,069.25
Vehicle	Central	0.20	0.10	0.08	2,537.15	3,234.78	2,232.52
	2	0.18	0.09	0.06	2,394.44	4,399.39	3,775.50
	3	0.19	0.11	0.09	3,878.74	4,435.79	3,269.74
	4	0.28	0.16	0.16	4,921.88	3,969.82	2,715.32
	5	0.21	0.08	0.07	3,667.30	4,567.11	3,883.59
	6	0.17	0.10	0.07	2,535.18	4,409.84	3,649.72
	7	0.20	0.09	0.08	3,152.53	4,488.47	3,733.92

Note: Blank cells denote the absence of recreational places in the respective isochrone.

Table 4. *Unique devices in each zone and the total GPS traces reported in the metropolitan area*

Zone	Name	Population	Total devices	Total GPS traces
Central	San Rafael	22,567	29,684	635,383
2	Chimalhuacán	24,542	4,471	111,843
3	M. Contreras	26,829	2,792	88,681
4	Milpa Alta	21,015	1,383	38,069
5	Chalco	18,422	1,805	50,970
6	Neza	21,099	4,561	135,022
7	Iztapalapa	18,043	1,899	68,392

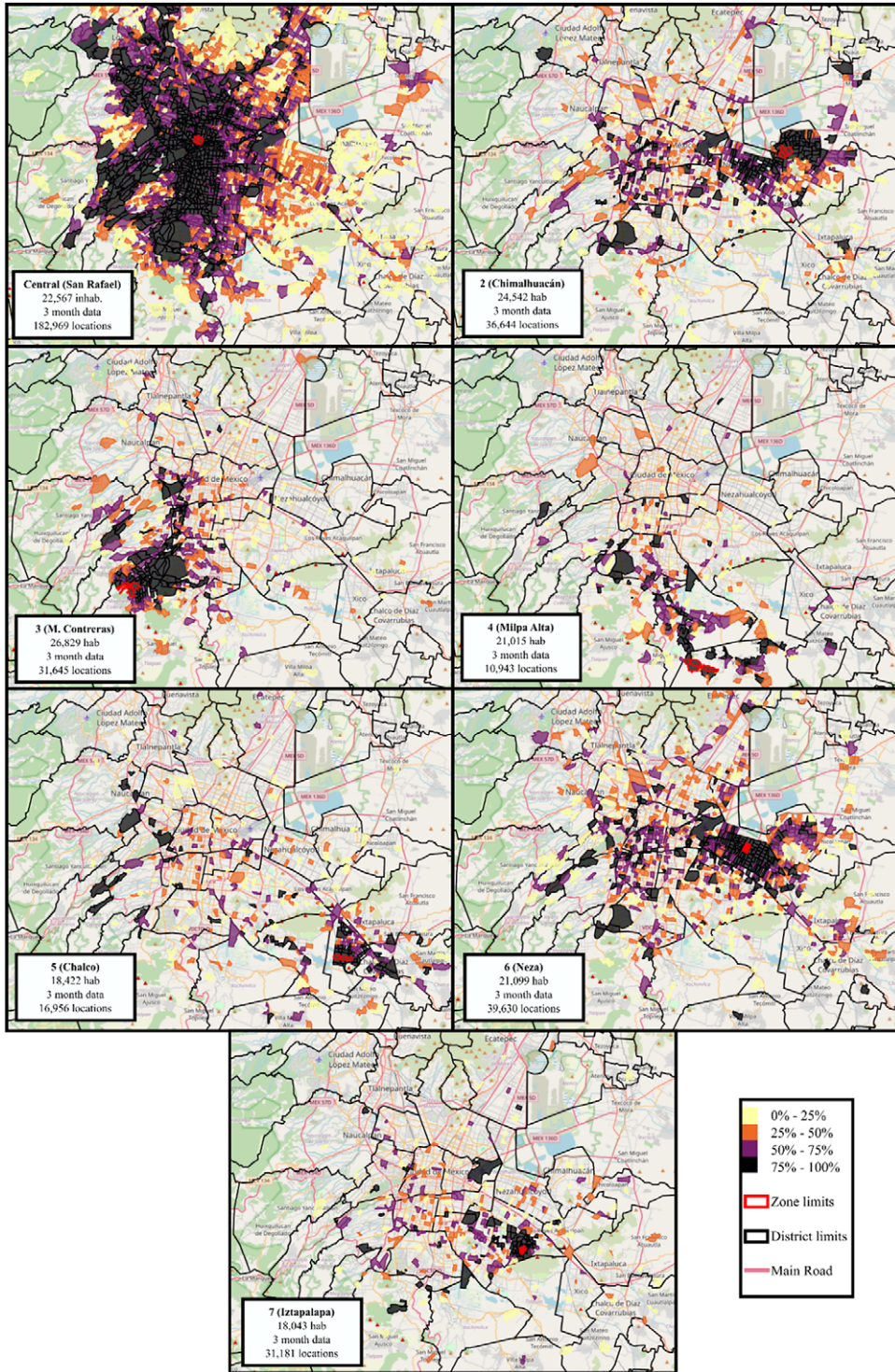


Figure 6. Origin-destination heat maps for each zone.

When comparing the maps, it has to be considered that the sample sizes are different, but the patterns are distinct. Visually, it is evident that home-based devices of the central Zone (San Rafael) have a much wider distribution of destinations. For the central zones devices, it is common that most of them are found closer and farther away in all cardinal directions. In contrast, the most common destinations for the suburban zones are near the home-AGEB. For all the suburban zones, the trips tend toward the central metropolitan area, where most workplaces and schools are located.

5.2. Distances from predominant AGEB

An indicator is developed to know how far or close the devices are found to their home-AGEB. The indicator is developed as follows. Centering in the home-AGEB are taken the following buffers:

- *Proximity* buffer: 2 km buffer centered in the home-AGEB.
- *Intermediate* buffer: The area between the 2 and 5 km buffer centered in the home-AGEB.
- *Distant* buffer: The area outside the 5 km buffer centered in the home-AGEB.

All the AGEBs contained in each buffer are selected if the device visits the AGEBs at least once in the 3 months. This selection allows computing the share of trips the devices make by distance to the home-AGEB.

To know the share of trips a device makes in their home-AGEB, we select all the device locations during the day (between 6:00 and 24:00) in the home-AGEB. As mentioned in Section 3, the home-AGEB is found by finding a high density of points between 1:00 and 6:00; the points in this interval are not considered because the device owner is probably sleeping. Table 5 presents the share of trips the devices make to each buffer and the trips within its home-AGEB.

Most device locations are within the home-AGEB, fluctuating notably between 62% and 89%. Interestingly, the central zone exhibits the lowest proportion, with only 62%. In contrast, Zones 2–6, with comparatively inferior private vehicle and transit accessibility, tend to have around 80% of devices in the home-AGEB. Further elaboration on this observation is provided later in the text.

The central zone portrays a notable share of both *proximity* (14%) and *distant* (17%) trips. As outlined in Section 4, this zone boasts a diverse array of leisure venues within a 15-min reach by foot, bike, transit, and private vehicle, potentially influencing short- and long-distance trips. Additionally, the central area enjoys robust transit connectivity, including subway lines, BRT, and buses. Approximately, 50% of households in this zone possess at least one vehicle, as detailed in Table 6 (Instituto Nacional de Estadística, Geografía e Informática, 2020).

The share of home-AGEB trips of Zones 2 and 5 is high. Transit access in these zones is only by franchise transit, which is unreliable and often poorly regulated. Also, these zones have the lowest share of car ownership in the studied zones, with an average of 27.5% of households owning at least one car. Poor

Table 5. Distance from home-AGEB distribution per zone

Zone	Name	Population	Collected locations	Home-AGEB	<i>Proximity</i> <2 km	<i>Intermediate</i> km < 5 km	<i>Distant</i> >5 km
Central	San Rafael	22,567	182,969	62%	14%	7%	17%
2	Chimalhuacán	24,542	36,644	86%	3%	5%	6%
3	M. Contreras	26,829	31,645	73%	8%	9%	10%
4	Milpa Alta	21,015	10,943	82%	4%	2%	12%
5	Chalco	18,422	16,956	89%	6%	1%	4%
6	Neza	21,099	39,630	79%	8%	5%	8%
7	Iztapalapa	18,043	31,181	66%	16%	13%	5%

Table 6. Share of households owning at least one vehicle according to the 2020 Census (*Instituto Nacional de Estadística, Geografía e Informática, 2020*)

Zone	Name	Share of households owning one or more than one vehicle
Central	San Rafael	50%
2	Chimalhuacán	28%
3	M. Contreras	40%
4	Milpa Alta	35%
5	Chalco	27%
6	Neza	38%
7	Iztapalapa	33%

transit connectivity and low car ownership could be related to why, in these zones, *distant* trips are not as common as in the other zones. However, further studies would be necessary. As seen in Section 4, Zone 2 has a good diversity and distribution of private entertainment places within walking distance from the centroid of Zone 2. However, the diversity and distribution of entertainment places in Zone 5 are low compared to the other zones. Therefore, further research is needed to explain why people in these zones prefer to stay home. Zone 6 is similar to Zones 2 and 5 regarding transit access and *distant* trips. However, this zone has one of the highest shares of car ownership among the suburban zones.

Zones 3 and 4 also have a high share of *distant* trips and the lowest share of local trips among all studied zones in the suburbs. On the one hand, in Zone 3, 40% of households have one or more cars. In contrast, in Zone 4, this share goes down to 35%, and both zones have access to franchise transit and buses operated by the local government. In general, in Zones 3 and 4, the transit accessibility is better than in the other studied suburban zones, so this could indicate why the share of *distant* trips is more significant in Zones 3 and 4 than in Zones 2, 5, and 6. Zone 3 has a regular-good mix and distribution of entertainment places (see Table 3) for the transit, bicycle, and pedestrian isochrones which can induce *proximity* trips and trips within the home-AGEB.

In Zone 7, 33% of households own at least one vehicle, and the transit access in this zone is good compared to the other studied suburban zones. This zone is close to two subway lines, one trolley line, one government-operated bus line, and franchise transit lines. However, in Zone 7, the share of *distant* trips is the lowest of all studied zones. It has the highest share of trips of *proximity* and *intermediate* trips. Zone 7 is close to the central metropolitan area and has a high share of *proximity* and *intermediate*. In contrast, the share of *distant* trips is considerably lower than the share of *proximity* of *intermediate* trips. In Zone 7, the mix and distribution of entertainment places for the transit, bicycle, and pedestrian isochrones are regular, that is, it is not as good as in Zone 2 or the central zone but is not as bad as in Zones 5 and 6. The low share of home-AGEB trips in Zone 7 could be partly explained by its closeness to the central metropolitan area, good access to private vehicles, transit, and a regular mix and distribution of entertainment places, making the people more prone to travel longer distances.

5.3. Device profiling

Three categories are introduced to profile the devices based on the duration of their trips. These categories aim to classify each device according to the length of its trips. For example, some devices may exhibit a significantly higher proportion of long trips compared to short trips. The profile classification is outlined as follows:

- Profile: *I don't travel much*. Comprising devices with more than 95% of their locations at the home-AGEB.

Table 7. *Percentage of people profiles per zone*

Zone	Name	<i>I don't travel much</i>	<i>Active close to home</i>	<i>I travel far</i>
Central	San Rafael	62%	21%	17%
2	Chimalhuacán	87%	4%	9%
3	M. Contreras	72%	13%	11%
4	Milpa Alta	79%	5%	16%
5	Chalco	77%	8%	5%
6	Neza	77%	11%	13%
7	Iztapalapa	78%	13%	9%

- Profile: *Active close to home*: Comprising devices with less than 80% of their locations at home-AGEB and more than 10% within a 2 km buffer.
- Profile: *I travel far*: Comprising devices with less than 80% of their locations at home-AGEB and more than 10% outside a 2 km buffer.

The identified profiles within this analysis encompass a portion of the sample, leaving 10.6% of individuals (devices) unclassified within these profiles. This 10% of individuals does not align with the defined patterns; they neither spend substantial time within their home-AGEB (between 80% and 95% of their locations at the home-AGEB) nor make frequent trips outside their home AGEB (less than 10% of their location within a 2 km buffer and outside a 2 km buffer). Some individuals (devices) fall into both the profiles of *Active close to home* and *I travel far*, suggesting simultaneous active engagement within their proximity and frequent visits to distant locations. It is observed that 1.9% of the sample falls into both categories.

For this analysis, devices with less than 10 reported locations were excluded. The profiling results are presented in [Table 7](#). The central zone exhibits the lowest share of devices in the *I don't travel much* profile (62%) and the highest shares in the *Active close to home* (21%) and *I travel far* profiles (17%). As mentioned earlier in the text, residents of the central zone enjoy good access to private vehicles and transit, and the zone boasts a well-distributed mix of entertainment venues across all isochrones.

In suburban zones, where access to private vehicles and transit is not as robust as in the central zone, and the mix and distribution of entertainment places are less favorable, the *I don't travel much* profile has a notably larger share than in the central zone.

Zone 4 exhibits poor transit connectivity, solely served by franchise transit, and features limited availability and distribution of public and private entertainment across pedestrian, bicycle, and transit isochrones. Moreover, it maintains a moderate level of car ownership compared to the studied zones. Notably, this zone reflects the second-largest share in the *I travel far* profile (16%), the lowest presence in the *Active close to home* profile (5%), and a significant share in the *I don't travel much* profile (79%).

5.4. *Walkability and local trips*

A walkability index based on the works of Gutiérrez-López et al. (2019), Institute for Transportation and Development Policy (2018), and Leticia et al. (2022) was estimated for all seven zones. The index considers the following parameters: urban grid, land mix, accessibility/proximity, distance between intersections, block density, accessibility to the public transport network, sidewalk width, accessible facades, access to green spaces, pedestrian crossings, urban canopy, slope, vehicular stream surface, lighting, pedestrian flow, and maximum speed.

The walkability index relies on a scoring system that incorporates the mentioned parameters. Each parameter is assigned a score between 1 and 3, with 1 indicating the lowest rank and 3 signifying the highest. These scores were derived from comprehensive data from field studies and GIS-based network

Table 8. Walkability index and share of Active close to home profile

Zone	Name	Walkability index	Active close to home profile
Central	San Rafael	39	21%
2	Chimalhuacán	15	4%
3	M. Contreras	23	13%
4	Milpa Alta	17	5%
5	Chalco	13	8%
6	Neza	17	11%
7	Iztapalapa	25	13%

analysis. After assigning scores to each parameter, the sum of these scores yields the walkability index for each zone. Detailed descriptions of the walkability parameters and the computations for the index can be found in the Supplementary Material.

Table 8 displays the walkability index for each zone along with the percentage of the *Active close to home* profile for that same zone. This profile pertains to devices that depart from their home-AGEB but remain within the 2 km buffer. Trips falling within the *Active close to home* profile are easily facilitated by active mobility and, as outlined by Pentland (2015), contribute to local engagement.

At first glance, the walkability of a zone is linked to the proportion of trips taken outside the home-AGEB but within a 2 km range. In fact, for these observations (central and suburban zones), there is a strong correlation ($R^2 = 0.815$) between walking conditions and the *Active close to home* profile. While this correlation holds for the zones examined in this study, its generalization and applicability to additional zones within the Metropolitan area remains untested but is a future area of research that should be explored, as these preliminary results are promising. Suppose the walking conditions correspond to a higher share of local trips, and it is assumed that these trips could boost local interactions. Then, improving the walking conditions of a zone could trigger local engagement.

6. Limitations

Acknowledging limitations is indeed critical for any study. It is essential to identify and address these limitations as they can affect the accuracy of the findings. This section will delve into specific limitations encountered in our study, providing a detailed discussion to contextualize their impact on the research outcomes.

In Section 3.1, the method used to identify home zones needs further refinement to enhance accuracy. There is a possibility of errors inherent in this method; however, quantifying these errors is challenging due to the anonymization of cell phone data. Nevertheless, it is worth noting that similar methodologies, as outlined in Section 3.1, are employed in other studies (Phithakkitnukoon et al., 2012; Jiang et al., 2017; Bayat et al., 2020) to identify home zones.

The absence of transit information introduces potential errors in computing the EI and HHI. Transit information would facilitate determining the precise reach of individuals using public transportation within a 15-min. While there is no formal study on transit waiting times for the Mexico City Metropolitan Area, media reports (Forbes, 2016) suggest an average wait time of approximately 11 min for any transit mode. Therefore, it is reasonable to assume that the 15-min transit isochrones closely resemble the pedestrian isochrones due to this similarity in timing.

Identifying cases where multiple individuals may have used a single device within our sample duration could involve detecting significant discrepancies in daily usage patterns; however, this type of analysis is outside the scope of the paper. Based on media reports (Statista Research Department, 2023), Mexicans typically replace their phones every 2 years and often pass down phones within households. Consequently, it is likely that during the 3-month study period, only a few devices were shared among multiple individuals across different households.

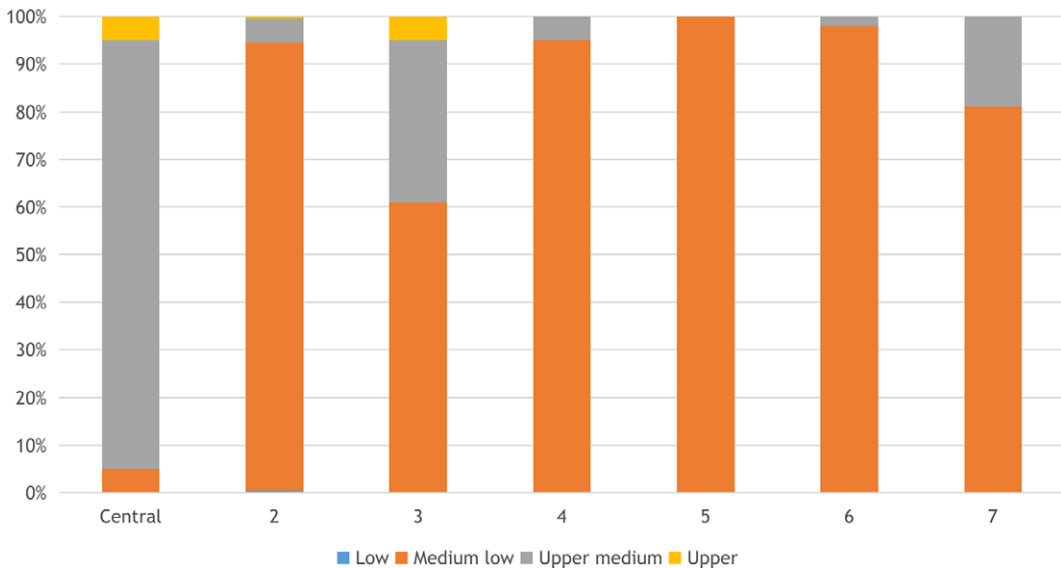


Figure 7. SES distribution in each zone.

The anonymized nature of the cell phone dataset limits the ability to determine the gender and age of the device user. The dataset only provides information regarding the device's OS. Consequently, conducting analyses based on gender or age becomes challenging. The most feasible approach to incorporate gender and age data would involve using an Origin-Destination survey and assuming that the share of trips by gender in the survey aligns with the cellphone dataset. However, this method could introduce substantial biases, for example, if a significant proportion of devices in the dataset are owned by males, leading to an incorrect gender assignment to the dataset. Given the potential for significant biases when analyzing trips by gender or age, this aspect falls outside the scope of this paper. Similar challenges would arise if attempting to categorize trips based on age groups.

The socioeconomic status (SES) in suburban zones, as shown in [Figure 7](#), predominantly leans toward medium–low categories. Zones 3 and 7 stand out with a significant share in the upper-medium sector. In contrast, the central zone showcases a predominantly upper-medium SES.

SES is not factored into this research for two main reasons. First, the dataset lacks associated SES information with each device. Additionally, to the best of our knowledge, data concerning cell phone ownership and SES specifically for the Mexico City metropolitan area is yet to be available. Thus, inferring the proportion of devices belonging to specific SES categories in the dataset becomes challenging. Similar to age and gender considerations, while the dataset might contain devices from a higher SES, the studied suburban zones primarily exhibit a medium–low SES. This discrepancy could introduce biases into the analysis.

Additionally, while some suburban zones display a notable share of short and long trips with diverse entertainment options, the correlation between trip length and SES might be challenging. The transportation accessibility in central and suburban zones differs significantly, potentially influencing trip lengths more than SES. Exploring if SES impacts short and long trips would require future research, possibly in a central (urban zone) zone with good accessibility and a high share of medium–low SES. However, this is beyond the scope of this study, which focuses on a descriptive analysis of trips in suburban zones of the Mexico City metropolitan area.

7. Conclusions and future directions

Mobile phone data holds significant potential for identifying issues, needs, and socio-economic characteristics within a given zone. In the central zone (San Rafael), there is a diverse range of entertainment

facilities, a pleasant walking environment and a high share of individuals falling into the *Active close to home* profile. Furthermore, the central zone benefits from robust transport connectivity, a high level of car ownership, and a notable number of devices falling into the *I travel far* profile. Hence, the central zone is presumably an ideal location for local engagement, providing residents with excellent opportunities for exploration. Zones 3, 6, and 7 (M. Contreras, Neza, and Iztapalapa) are similar to the central zone as these zones feature a good distribution and diversity of entertainment services, coupled with the highest walkability indexes among the studied suburban zones and a high share of trips in the *Active close to home* and the *I travel far* profiles. While transit access is not as good in these zones as in the central zones, these zones have the best transit access among the studied suburban zones.

Zones 2, 4, and 5 (Chimalhuacan, Milpa Alta, and Chalco) have poor transit coverage and exhibit the highest proportion of individuals in the *I don't travel much* profile, suggesting a need for increased opportunities for exploration among residents in these areas. More transit alternatives are deemed necessary for these zones. Zones 2 and 4 demonstrate the lowest share of the *Active close to home* profile. Moreover, their distribution and diversity of entertainment places fall short compared to other zones across all isochrones. Zones 2 and 4 could significantly benefit from enhanced transit coverage and government intervention in public spaces that could foster better engagement and exploration opportunities for their residents.

Despite its limited connectivity and lower income level, Zone 4 boasts the highest proportion of individuals in the *I travel far* profile compared to other suburban zones studied. Zone 4 performs poorly on the diversity and distribution of entertainment services accessible within pedestrian, cycling, and transit isochrones. The high proportion of long trips and the absence of entertainment services could indicate that residents of Zone 4 need to allocate time and resources for extensive journeys in search of entertainment. Consequently, Zone 4 presents an opportune location for investment in local infrastructure and services that could foster local engagement.

The descriptive analysis of mobile data offers many potential applications, particularly when combined with other statistical and descriptive analyses of the urban landscape. In the context of public space, the study suggests that public intervention in walking conditions and entertainment venues might be related to local trips, which could increase community interactions and strengthen local dynamics. So, it might be reasonable for governments to invest in walking conditions and public entertainment to improve people's quality of life.

This research holds the potential to guide interventions tailored to specific population needs. Analyzing GPS traces from mobile phones helps devise indicators that gauge residents' proclivity to travel in different areas. Profiling inhabitants based on their mobile phone GPS traces sheds light on whether they confine their activities to their zone or frequently engage in short and long trips.

The methods presented in this study can be customized for analyzing GPS traces from diverse regions. The insights derived from this descriptive analysis can be invaluable for stakeholders and government decision-makers in formulating land use and accessibility strategies. Policymakers seeking to propose services or modify land use policies at metropolitan levels can base their decisions on studies like the one presented here. Moreover, this study's framework can be adapted or replicated for other pertinent projects reliant on verifiable geographical data. Acquiring GPS traces from mobile phones not only reduces implementation costs but also has the potential to identify mobility problems in specific neighborhoods.

In future studies, examining the impact of specific transportation projects or temporary closures of transportation services on the population would be valuable. Conducting a before–after analysis, akin to the one presented in Sections 5.2 and 5.3, could provide insights into how the closure or reopening of transportation services affects trip distribution, offering invaluable assessments of the benefits arising from new transportation services.

Furthermore, future research aims to expand the analysis by considering additional zones within the central and suburban areas of the Mexico City metropolitan region. This broader analysis will contribute to a more comprehensive understanding of trip profiles within the metropolitan area. Additionally, in characterizing short trips, it would be worthwhile to consider the diversity and distribution of other services, such as schools, markets, and convenience stores.

Exploring the correlation between gentrification and the attraction of the floating population during the day would be an intriguing case study that could be effectively examined using GPS traces. In this scenario, the initial task would involve identifying all mobile IDs within a zone during the daytime and then subtracting those belonging to the local zone. Subsequent analysis would focus on understanding the origin, frequency, and distance of external devices encountered and whether they can be correlated to gentrification indicators.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/dap.2024.8>.

Data availability statement. Replication data found in <https://osf.io/beh4j/>.

Acknowledgments. The authors of this paper would like to thank *Dat's Why* for providing the mobile phone data for this research. The authors would also like to thank Jan Wegner, Katherine Hay, Luisa Rubio, and Daniel Rodriguez for their valuable comments on improving this paper.

Author contribution. Conceptualization: A.S., D.L., C.D., O.R.; Data curation: O.R.; Formal analysis: A.S.; Funding acquisition: A.S., D.L., C.D., O.R.; Investigation: A.S., D.L., C.D.; Methodology: A.S., D.L., C.D., O.R.; Project administration: A.S.; Visualization: D.L.; Writing—original draft: A.S., D.L.; Writing—review and editing: A.S., D.L., C.D., O.R.

Funding statement. This research was supported by grants from the CAF—Banco de Desarrollo de América Latina; TUMI—Transformative Urban Mobility Initiative. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interest. The authors declare no competing interests exist.

Ethics statement. The research meets all ethical guidelines, including adherence to the legal requirements of the study country.

References

- Badawi S and Farag AA** (2021) Young Saudi women's travel behavior change over 2015/2020. *Journal of Transport & Health* 21, 101080.
- Bayat S, Naglie G, Rapoport MJ, Stasiulis E, Chikhaoui B and Mihailidis A** (2020) Inferring destinations and activity types of older adults from GPS data: Algorithm development and validation. *JMIR Aging* 3(2), e18008.
- Chand S, Yee E, Alsultan A and Dixit VV** (2021) A descriptive analysis on the impact of COVID-19 lockdowns on road traffic incidents in Sydney, Australia. *International Journal of Environmental Research and Public Health* 18(21), 11701.
- Eagle N, Pentland AS and Lazer D** (2008) Mobile phone data for inferring social network structure. In Liu H, Salerno JJ and Young MJ (eds), *Social Computing, Behavioral Modeling, and Prediction*. Boston, MA: Springer US, pp. 79–88.
- Eagle N, Pentland AS and Lazer D** (2009) Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences* 106(36), 15274–15278.
- Ebrahimpour Z, Wan W, García JLV, Cervantes O and Hou L** (2020) Analyzing social-geographic human mobility patterns using large-scale social media data. *ISPRS International Journal of Geo-Information* 9(2), 125.
- Forbes** (2016) Mexicanos pierden hora y media diario en el transporte público. <https://www.forbes.com.mx/mexicanos-pierden-hora-media-al-dia-transporte-publico/> (accessed February 2023).
- Guo S, Song C, Pei T, Liu Y, Ma T, Du Y, Chen J, Fan Z, Tang X, Peng Y and Wang Y** (2019) Accessibility to urban parks for elderly residents: Perspectives from mobile phone data. *Landscape and Urban Planning* 191, 103642.
- Gutiérrez-López JA, Caballero-Pérez YB and Escamilla-Triana RA** (2019) Índice de caminabilidad para la ciudad de Bogotá. *Revista de Arquitectura* 21(1), 8–20.
- Heikinheimo V, Tenkanen H, Bergroth C, Järv O, Hiipala T and Toivonen T** (2020) Understanding the use of urban green spaces from user-generated geographic information. *Landscape and Urban Planning* 201, 103845.
- Institute for Transportation and Development Policy** (2018) *Pedestrians First, Tools for a Walkable City*, 1st edition. New York: Institute for Transportation and Development Policy.
- Instituto Nacional de Estadística, Geografía e Informática** (2020) *Censo de Población y Vivienda 2020*. Aguascalientes, Mexico: INEGI. <https://www.inegi.org.mx/programas/ccpv/2020/> (accessed February 2023)
- Instituto Nacional de Estadística, Geografía e Informática** (2023) *Directorio Estadístico Nacional de Unidades Económicas*. Aguascalientes, Mexico: INEGI. <https://www.inegi.org.mx/app/mapa/denue/default.aspx> (accessed February 2023)
- Jacobs J** (1992) *The Death and Life of Great American Cities*. New York: Vintage Books.
- Jiang S, Ferreira J and Gonzalez MC** (2017) Activity-based human mobility patterns inferred from mobile phone data: A case study of Singapore. *IEEE Transactions on Big Data* 3(2), 208–219.
- Katrakazas C, Michelaraki E, Sekadakis M and Yannis G** (2020) A descriptive analysis of the effect of the COVID-19 pandemic on driving behavior and road safety. *Transportation Research Interdisciplinary Perspectives* 7, 100186.

- Leticia S, Tini B, Sato B, Farias D and Pitombo F** (2022) *Metodología para calcular el índice técnico de caminabilidad sensible al género*. Technical Report 2530, Banco Interamericano de Desarrollo (BID).
- Pentland A** (2015) *Social Physics: How Social Networks Can Make Us Smarter*. New York: Penguin Press.
- Phithakkitnukoon S, Smoreda Z and Olivier P** (2012) Socio-geography of human mobility: A study using longitudinal mobile phone data. *PLoS One* 7(6), e39253.
- Reif J and Schmücker D** (2020) Exploring new ways of visitor tracking using big data sources: Opportunities and limits of passive mobile data for tourism. *Journal of Destination Marketing & Management* 18, 100481.
- Shi L, Chi G, Liu X and Liu Y** (2015) Human mobility patterns in different communities: A mobile phone data-based social network approach. *Annals of GIS* 21(1), 15–26.
- Sila-Nowicka K, Vandrol J, Oshan T, Long JA, Demšar U and Fotheringham AS** (2015) Analysis of human mobility patterns from GPS trajectories and contextual information. *International Journal of Geographical Information Science* 30(5), 881–906.
- Song Y, Merlin L and Rodriguez D** (2013) Comparing measures of urban land use mix. *Computers, Environment and Urban Systems* 42, 1–13.
- StatCounter** (2023) *Mobile Operating System Market Share Mexico*. Dublin, Ireland: StatCounter. <https://gs.statcounter.com/os-market-share/mobile/mexico/#monthly-201901-201912> (accessed November 2023).
- Statista Research Department** (2023) *Frecuencia del cambio de teléfonos móviles de los usuarios en México en 2020*. Hamburg, Germany: Statista. <https://es.statista.com/estadisticas/1233037/frecuencia-cambio-telefonos-moviles-usuarios-mexico/> (accessed March 2023).
- Takaki E** (2018) Mobile technology (gps) and spatial appropriation in public spaces in São Paulo, Brazil. In *ISUFItaly Bari 2018*, pp. 551–560. Bari, Italy.
- TravelTime** (2023) *Build Brilliant Apps with TravelTime Location APIs*. London, United Kingdom: ITO World Limited. <https://traveltime.com/> (accessed March 2023).
- Valhalla** (2023) *Valhalla Routing Service API Reference*. Pennsylvania, USA: Mapzen. <https://github.com/valhalla> (accessed March 2023).
- Wang Z, He SY and Leung Y** (2018) Applying mobile phone data to travel behaviour research: A literature review. *Travel Behaviour and Society* 11, 141–155.
- Xiao Y, Wang D and Fang J** (2019) Exploring the disparities in park access through mobile phone data: Evidence from Shanghai, China. *Landscape and Urban Planning* 181, 80–91.
- Zhao X, Lu X, Liu Y, Lin J and An J** (2018) Tourist movement patterns understanding from the perspective of travel party size using mobile tracking data: A case study of Xian, China. *Tourism Management* 69, 368–383.
- Zou Z, Younes H, Erdoğan S and Wu J** (2020) Exploratory analysis of real-time e-scooter trip data in Washington, D.C. *Transportation Research Record: Journal of the Transportation Research Board* 2674(8), 285–299.