Street vending is a complex systemic phenomenon in most cities worldwide, with different intensities and features. In the Mexican case, it is an activity with remnants of a precolonial logic in its spatial distribution. Thus, a low correlation exists between the street vending government regulations and the actual day by day organization of the activity. Certain authors have suggested and compiled an econometric model that considers some variables to comprehend the phenomenon better. All the variables came with the detailed information except for the territorial aspect. Thus, an accessibility tool was created to provide a robust location profile, using official variables related to socioeconomic topics recommended by the World Bank. The resulting database was then analyzed by Machine Learning prediction models. The results provided a map with spatial categorization of the street vending activity, with a solid correlation (0.509 ± 0.039) to the jobs variable.

1. Introduction

Street vending (SV) activity is as old as the story of the cities (Contreras & Weiht, 1988), partly because there is an essential part of commercial culture in the origin and subsequent survival of societies, hence, the cities. Therefore, it represents an essential element in practically all commercial cities. Other researchers exposed the paintings of thousands of years ago in Roman and Chinese empires that depict this phenomenon (Sun et al., 2016), stating the difficulty in administering and controlling this activity because of a self-organizing tendency. Today, this implies a risk relationship with any city government.

The recent situation of SARS-CoV-2 has a double factor. There are high estimations of job losses, which means more people potentially adopting the activity (ILO, International Labour Organization, 2020). On the other hand, there is a positive change in the process of recognizing the street as a habitat that should help provide space for open-air activities (NACTO, National Association of City Transportation Officials, 2021). Considering the street network of a city as public space and also as a habitat requires an innovative tool to measure it. Public space accessibility gives a new meaning and sustainable benefits to the streets, beyond the known implications of walking, cycling, or public transport mobility versus moving in private and motorized transport (Jacoby & Pardo, 2012; Pivo & Fisher, 2011), including the habitability variables, such as proximity to schools, jobs, commerce and others, implying the importance of the location variable. The latter is important since a recent econometric model made in Portland by Glicker (2014), considers that the spatial aspect has a significant impact on the ability of the street vendors to succeed. It is important to remember, that econometric models are created to better comprehend a phenomenon (Bjerkholt, 2016).

However, Mexican research in this aspect has the spatial variable not represented enough. It is only mentioned in the Domínguez-Gijón and Venegas-Martínez (2016) study, not as a variable but as a cost variable not represented enough. It is only mentioned in the Domínguez-Gijón and Venegas-Martínez (2016) study, not as a variable but as a cost of government permission to sell (Domínguez-Gijón & Venegas-Martínez, 2016). Meanwhile, García-Guzmán (2001) reaches a categorization of two large groups: legal or illegal zones for SV activity (García-Guzmán, 2001). This is because of the negation of this millennial activity by some social sectors and even official road design. It implies the necessity of a paradigm change to improve the conditions of the people involved (Crossa, 2017). It is important to remember that this negation came from hundreds of years ago. The Second Relation Letter by Cortes in 1520, reveals a priority use in the Aztec logic in spatial
distribution, creating a diversity of plazas that gathered around 20,000 people on regular days and up to 60,000 on market days (Thatcher, 1907). The New Spain government imposed a spatial process, limiting the activity to a certain population number, inside a market building. This started a recurrent overpopulation until there were more vendors outside than inside. This led to the self-organization based on government regulation limits that predate the new millennia (García-Guzmán, 2001).

The proposal of this investigation reinforces the spatial characteristic of the activity, creating a Main Map with official information, that categorizes street vending activity. For this, we use a tool that needs to be explained, with regard to considering a street network as public space. The process considers three steps: First in Section 2, we provide some necessary information about the activity in a selected metropolitan area: ZM Monterrey. Then, in Section 3, we propose construction of a Public Space Accessibility Tool (PSAT) helping to characterize the SV activity based on spatial geolocated variables. Finally, in Section 4, the resulting database was analyzed, incorporating prediction models extracted from Machine Learning, ‘Boosted Regression Trees’ (BRT), Deep Learning, and clustering to complete the categorization of certain ecosystems for SV activity.

2. Materials: basic information of street vending activity

2.1. Street vending in Mexican cities

The Mexican Constitution recognizes the right to work, and it addresses disputed negotiations in SV organizations and municipalities. Meneses-Reyes (2018) reports a constant dispute over the space, acknowledging that the street is not only for human or product transit but a constitutional guarantee of freedom to work and commerce that is in constant limitation by authority day by day (Meneses-Reyes, 2018). Silva Londoño (2010) explains a process in which SV activity suffers from pressure by authorities who try to inhibit the organization of the vendors, sometimes using police provocation (Silva Londoño, 2010). Another perspective is the notion of legality according to the general view of society. Hart (1973) defined legitimate activities as those that contribute to economic growth at a small scale, such as home-based production, manual labor, or personal services. Illegitimate informal activities are not necessarily criminal but are of questionable value to national development (Hart, 1973).

2.2. Street vending localization

Recently, official and georeferenced information was made available by INEGI (Instituto Nacional de Estadística y Geografía) to better appreciate the relevance of this activity. There are two large SV official categorization groups by the Encuesta Nacional de Ocupación y Empleo (ENOE): fixed and semi-fixed trade. This is mapped with points of every location, adding statistics such as the number of employees and type of activity. The exercise is performed with the semi-fixed (SV-SF) trade for practical purposes. Semi-fixed includes every person or group that conducted commercial activity on the street, using structures, vehicles, or any kind of furniture installed temporarily every day of the activity, being removed once the day is over. This activity has a better regulation characteristic than other types of street vending.

The Monterrey metropolitan area has a considerable size. Therefore, the data were converted into a grid as Krambeck did in 2006 for representative surveys source (Krambeck, 2006), and the Salvador Rueda grid technique in the Index Plan for Vitoria Gasteiz (Rueda, 2010). Recently, it was used by Pánek (2019) for a participatory planning support system in Olomouc, Czech Republic (Pánek, 2019). The grid is hexagonal with a size approximately of 400 m from the center to the sides, hence it emulates a five-minute walking distance circle. The grid measures the density ranges of street vending activity, summing all activity points inside each hexagon, using Quantum Geographic Information Systems (QGIS) software. As Figure 1 shows, the SV-SF activity in the Monterrey metropolitan area happens across the municipalities.

2.3. Influence of street vending

Recent innovative investigation contemplates the SV activity as having the potential to do more good than harm to a city, both the spaces and the inhabitants (Ehrenfeucht, 2017; Lewis & Draeger, 2011; Newman & Burnett, 2013; Rodgers & Roy, 2010; Sun et al., 2016). Sun et al. (2016) concluded that street vending as a public space activity has enough characteristics to be part of and create a special place in an actor-network theory. They relate that an SV network can be a socioeconomic structure in which this kind of relationship appears. Meanwhile, Ehrenfeucht (2017) worked on the possible intersections regarding the pedestrians, brick and cement trade locals, and street vending activity, using observation techniques and interviews. She concluded that 80 % of trade locals had a favorable opinion of the activity. Meanwhile, Rodgers and Roy (2010) attributed the SV to part of the tourist attraction and city personality of Portland. Lewis and Draeger (2011) talked about a warm glow effect in that ‘anyone gets to enjoy, for minutes at least, the feeling of being part of a larger community’. Finally, Newman and Burnett (2013) researched street food characteristics in Portland. They concluded there
are two essential characteristics to understand the success: a progressive and social normative and the urban context quality (Newman & Burnett, 2013).

The Mexican case has some studies on SV and the public space (Herrera, 2017; Silva Londoño, 2010). Silva Londoño (2010) suggested a network analysis to comprehend a particular urban space and how SV activity capacity allows creation and formulation of a social and physical space articulation. Meanwhile, Verónica Crossa (2017) added to the need to acknowledge the street as a public space. She expresses that ‘public space is used as a neutral and apolitical concept, but the street does not…’, adding that a technical justification against mobility projects is really just a cover for unfair criteria (Crossa, 2017).

3. Method: Phase A. Construction of a public space accessibility tool for characterizing street vending activity

3.1. Accessibility review

Accessibility is a form of measuring the physical nearness or proximity between a person or a thing concerning another person(s) or thing(s) in a defined area. There is still a necessity for a unified theory for the accessibility phenomenon to help in the modeling and process of urban mobility and transport planning (Batty, 2009). Since the 1900s, there have been some techniques and methodologies for measuring this concept, most of them evolving with new technologies (Pirie, 1979). One measure used is cumulative opportunities, an exercise in which the sum of activities are valued, adding a time or distance cost, given with an isochrones zoning based on an origin.

3.2. Construction of public space accessibility tool

The construction of a systemic mathematical model for analyzing a city, needs at least two essential elements: the variables and the spatial or zoning area (Haggett & Chorley, 1967). In this case, the zoning area is the urbanization limits of the Monterrey metropolitan area. For the variables, the street as a public space network produces a criteria base of variables. There have been efforts to create an index to measure these public space activities in the last two decades. First, the walkability index, an effort of a study by the Massachusetts Institute of Technology and funded by the World Bank (Krambeck, 2006), helps provide a technological tool for developing countries. In 2007, the Walk Score, a more fundamental and free instrument was created to measure the accessibility of walking by the sum of all opportunities around a place (Score, 2014). This online application uses a cumulative opportunities technique, and also adds habitability variables such as proximity to bars, gyms, and other amenities. Then, some authors confirmed the value
of the index, such as Pivo and Fisher (2011), who reported that high levels of a walk score represent a higher value to the price of specific building segments (Pivo & Fisher, 2011). Salvador Rueda Index Plan for Vitoria-Gasteiz Rueda (2010) include an indicator for the habitability of public space and services, and it included some variables as proximity to schools, jobs and commerce among others (Rueda, 2010). International consultancy agencies retook the concept (ITDP, Institute for Transportation and Development Policy, 2018), validating the variables with a two-scale approach. Later, studies by Sofwan and Tanjung (2020) and Carpio-Pinedo et al. (2021) evolved the technological process, but the main variables did not change. Their results indicate the potential to predict non-motorized activity in the streets. As Ortega et al. (2020) admits, technological and information availability limits the universe of certain variables; meanwhile, others expand them (Ortega et al., 2020).

Sofwan and Tanjung (2020) review of the innovation process during the last 20 years of walkability indices elaborates the basic idea by Krambeck (2006) of adding value to the mixed-use variables, hence the cumulative opportunities logic.

3.3. Selection of variables and values

Given how the variables are presented in official geo-referenced data, the result is a total of 14 variables grouped into three dimensions: Infrastructure quality, Attraction, and Structure, see Table 1. The majority of them are from the original Krambeck (2006) catalog, adding the habitability variables from the Rueda (2010) proposal. The availability of the variables were confirmed with official and geo-referenced information by INEGI (INEGI, Instituto Nacional de Estadística y Geografía, 2010, 2018, 2019).

All variables obtain a value to all streets segments of the Monterrey metropolitan street network, as follows:

(a) **Infrastructure quality.** We value the street segments based on the existence of sidewalks, signalization, street lights, accessibility ramps, and trees, with information available from official database (INEGI, Instituto Nacional de Estadística y Geografía, 2018).

(b) **Attraction.** We valued the same street segments, with a level of attraction based on activities that exist in the street. We weigh the activities by the distribution of the travel motivations: jobs, schools (preferably public), and local commerce (corner shops and small business), and the return to home (housing). The information was provided from official data base (INEGI, Instituto Nacional de Estadística y Geografía, 2019).

(c) **Structure.** This was valued by the road hierarchy of the street using the Manual de Calles, diseño vial para ciudades mexicanas (SEDATU, Secretaria de Desarrollo Agrario, Territorial y Urbano, 2018), as well as the interconnectivity with all forms of public transport (Monterrey has Metro and BRT systems), longitude of the blocks (around 100 m), and a special polygon such as the City Center or special districts. The exercise took out all the streets with bad slope percentage (above of 14 percent). Furthermore, it took out all streets from closed neighbors. The information was provided by official data base (INEGI, Instituto Nacional de Estadística y Geografía, 2019) and by OpenStreetMap (OSM).

The three dimensions had a value of 100 points each, with a first review consisting of exercises and evaluation by 50 interdisciplinary students of architecture and geography that qualified every variable inside the dimensions, hence totaling the sum of 100. A second review, made using focus groups with academics, activists and specialists of the Monterrey metropolitan zone, asking to qualify and differentiate the value of each variable. The final values are reflected in Table 2.

3.4. Database

With the latter results, the next step was to run the algorithm in the qGIS, valuing each of the nearly 95,000 street segments in the metropolitan road network. The value used for this exercise was made with the first review, given the availability of this one. Then, in Figure 2, we converted the data to the same reticular grid of hexagons with a center-to-side distance of 400 m, enabling a macro-scale geospatial

| Table 1. Selected variables for analyzing public space accessibility. |
|---------------------------------------------------------------|
| Sidewalks | Proximity to schools | Road hierarchy |
| Ramps    | Proximity to housing | Transport existence |
| Trees    | Proximity to jobs    | Longitude of blocks |
| Signals  | Proximity to commerce| Slope |
| Street lights | Proximity to commerce | Special polygon |

| Table 2. Quantification process of 13 selected variables for the evaluation. |
|-----------------------------------------------|
| Dimensions | Variables | First review | Focus groups |
|-------------|-----------|--------------|--------------|
| Infrastructure | Sidewalks  | 40            | 35            |
|             | Ramps     | 20            | 10            |
|             | Street lights | 10          | 22            |
|             | Signals   | 10            | 8             |
| Attraction  | Proximity to jobs | 40            | 32            |
|             | Proximity to schools | 25           | 30            |
|             | Proximity to housing | 20           | 25            |
|             | Proximity to commerce | 15           | 13            |
| Structure   | Longitude of blocks | 10           | 10            |
|             | Road hierarchy  | 10            | 15            |
|             | Public transport existence | 40           | 40            |
analysis but also related to a scale of a pedestrian network.

The SV-SF population in the Monterrey metropolitan area consists of 3850 sites, with diverse services such as traditional Mexican food with tacos, tostadas, gorditas, and foreign food such as hamburgers, pizzas, and hot dogs. Then, there are non-food services such as shoeshiners, handicrafts, cosmetics, bazaars, jewelry, and locksmiths. With this database, we proceeded to the Machine Learning analysis with BRT, deep learning, and clustering.

4. Method: Phase B. Machine learning process for categorization of street vending activity

For the classification problem, the techniques used were Deep Learning, gradient boosted tree, and clustering.

4.1. Deep learning

To establish a prediction based on the data, Deep Learning uses a multilayer feed-forward artificial neural network that is trained with a stochastic gradient descent using backpropagation. The network can contain a large number of hidden layers consisting of neurons with tanh, rectifier, and maxout activation functions. Advanced features such as adaptive learning rate, rate annealing, momentum training, dropout, L1 or L2 regularization, check-pointing, and grid search enable high-predictive accuracy. Each node trains a copy of the global model parameters on its local data with multi-threading (asynchronously) and contributes periodically to the global model via model averaging across the network. The correlation in the category of work in terms of the other categories gave 0.509 ± 0.039, an acceptable but not excellent result.

4.2. Boosted gradient tree

Gradient boosting is a technique for regression analysis and statistical classification problems, which produces a predictive model in the form of a set of weak prediction models, typically decision trees. It builds the model in a staggered way as other boosting methods do, and generalizes them by allowing arbitrary optimization of a differentiable loss function for the classification of the data categories. The selected categories for the present work are food, restaurants, trade, and service. From previous research in Figure 3, the predictions for the research were generated in terms of the classification of job categories: signals, trees, street lights, schools, commerce, housing, road hierarchy, public transport, schools, and ramps. The activities were in turn divided into services, food, fast food and, trade (cosmetics, jewelry, haberdashery). The result showed that there was a good correlation between inputs and the intended forecast. It can be seen that there is a relationship between work and the rest of the categories such as trade, merchants,
services, fast food, and access, in terms of the job category.

4.3. Clustering

We decided to employ a clustering technique to aid in the visual identification of groups of roads sharing characteristics based on the PSAT score. The analysis led to nine clusters. The nine clusters were described based on the composition of the businesses on the roads in order to identify significant elements that could encapsulate them. The details of the analysis are described below.

Because the tree-based classifier failed to outstandingly classify the 30 types of establishments located in the city based on PSAT, a cluster analysis was conducted to explore how blocks are grouped based on their PSAT score.

One of the elements necessary to perform a cluster analysis is to provide, a priori, the number of groups that one believes may be present. This number can be determined by various methods ranging from a visual analysis to statistical tests such as the likelihood ratio (see Hardy, 1996). In that sense, it was decided to perform a visual exploration using t-SNE (t-distributed Stochastic Neighbor Embedding). t-SNE enabled visualizing in two dimensions the possible clusters in a dataset; at the same time, it is considered a clustering method (Linderman & Steinerberger, 2019). However, the dimensions generated by this method are not as interpretable as they are, in many cases, when a dimensionality reduction method such as Principal

Figure 3. Boosted gradient tree.

Figure 4. t-distributed stochastic neighbor embedding (t-SNE) of PSAT after eliminating duplicates and using perplexity configuration of 30.
Component Analysis (PCA) is applied (see Wattenberg et al., 2016). Figure 4 clearly shows nine well-defined clusters by the Partitioning Around Medoids (PAM) algorithm on the data embedding generated by the t-SNE algorithm. When analyzing the composition of each of the 9 clusters, it turns out that food stores, general stores, coffee shops, general restaurants, and automotive service stores were present in most of the nine clusters, while jewelry, notions stores, and clothing stores were concentrated in a more limited number of groups (Table 3). The representation was obtained using the M3C package of R (see John et al., 2020). Clustering was obtained using the PAM algorithm on the embedding. In this sense, it seems that the decision on the location of a establishment is independent of the characteristics of the blocks, especially restaurants. This could indicate that the business plan on feasibility does not consider the PSAT score of the block to determine the location of the establishment, either because it is irrelevant or because it is ignored that it could be relevant.

Furthermore, looking for the possibility that the Simpson’s Paradox might be present, separability was analyzed using more general labels: Retail, Food, and Services (Figure 5). However, in this case, the paradox was not present, as the groups determined by the 30 labels were equally nonseparable as the groups formed by the labels: Retail, Food, and Services.

In the case of blocks, it is common for there to be several stores on the same block. This explains the inseparability of the labels. However, it could be that certain clusters have a higher concentration of one kind of establishment than another kind. Table 4 shows the distribution of establishments using the labels: retail, food, and services by clustering each group identified by the t-SNE algorithm. The Chi-square test was performed to determine the independence between the t-SNE clusters and the three establishment types mentioned above. The result was that they are not independent, since the chi-square statistic

### Table 3. Comparison between original classes and clusters.

| Establishment type/Cluster | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  |
|----------------------------|----|----|----|----|----|----|----|----|----|
| Grocery store              | 3  | 0  | 5  | 4  | 1  | 4  | 8  | 1  |    |
| Bazaar                     | 0  | 0  | 5  | 2  | 0  | 5  | 1  | 4  |    |
| Trinket store              | 0  | 0  | 0  | 1  | 0  | 0  | 1  | 1  |    |
| Cellphone store            | 1  | 0  | 3  | 4  | 0  | 0  | 1  | 0  |    |
| Makeup store               | 7  | 1  | 4  | 6  | 0  | 1  | 0  | 1  |    |
| Food store                 | 2  | 11 | 1  | 6  | 0  | 5  | 1  | 4  |    |
| Jewelry store              | 2  | 0  | 5  | 9  | 0  | 0  | 7  | 0  |    |
| Notions store              | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 1  |    |
| Other stores               | 15 | 4  | 12 | 16 | 4  | 9  | 16 | 7  | 6  |
| Clothing store             | 0  | 0  | 0  | 1  | 0  | 4  | 0  | 2  |    |
| Lottery kiosk              | 1  | 1  | 1  | 1  | 1  | 2  | 0  | 1  |    |
| Cafeteria                  | 22 | 4  | 52 | 7  | 10 | 15 | 10 | 12 | 4  |
| Fast food restaurant       | 1  | 0  | 2  | 1  | 1  | 7  | 1  | 3  |    |
| Fast Mexican food restaurant | 3  | 7  | 11 | 3  | 7  | 13 | 4  | 5  | 4  |
| BBQ restaurant             | 1  | 0  | 1  | 1  | 1  | 2  | 0  | 1  |    |
| Fonda                      | 2  | 3  | 3  | 6  | 1  | 4  | 4  | 1  | 2  |
| Gorditas restaurant        | 3  | 4  | 5  | 2  | 4  | 4  | 3  | 8  | 5  |
| Hamburger restaurant       | 4  | 5  | 5  | 5  | 6  | 15 | 5  | 9  | 2  |
| Seafood restaurant         | 2  | 0  | 0  | 1  | 0  | 1  | 3  | 2  | 1  |
| Chicken restaurant         | 3  | 4  | 12 | 4  | 3  | 8  | 11 | 8  | 4  |
| Taqueria                   | 38 | 42 | 73 | 52 | 41 | 95 | 42 | 60 | 35 |
| Sandwich shop              | 1  | 3  | 1  | 3  | 2  | 1  | 1  | 0  | 0  |
| Tostadas restaurant        | 3  | 3  | 9  | 0  | 2  | 15 | 4  | 4  | 3  |
| Food cart                  | 1  | 3  | 3  | 10 | 9  | 3  | 1  | 2  |    |
| Automotive service         | 1  | 5  | 2  | 3  | 3  | 7  | 2  | 4  | 3  |
| Shoe shiner service        | 5  | 2  | 5  | 10 | 0  | 3  | 2  | 0  | 3  |
| Locksmith’s shop           | 1  | 0  | 0  | 3  | 1  | 5  | 3  | 0  | 0  |
| Beauty salon               | 1  | 1  | 0  | 1  | 0  | 3  | 0  | 2  | 0  |
| Manufacturing service      | 2  | 0  | 2  | 2  | 1  | 3  | 1  | 2  | 2  |
| Other Services             | 5  | 2  | 13 | 16 | 3  | 10 | 5  | 5  | 1  |

![Figure 5](image-url)
obtained was 48.214 and the associated $p$-value was .0000439.

5. Results

To further explore the visual findings provided by the t-SNE technique, the PCA technique was used. PCA is also a visualization technique; however, in this case it was used to understand more directly the clusterings found using t-SNE. In this way, we were able to describe the groups directly using the indicators that make up the PSAT.

Most of the clusters are composed of between 70% and 80% food establishments. However, cluster 3 deviates from this trend because it only has a 51.50% concentration of this kind of establishments. Since the t-SNE dimensions were not interpreted, the PCA algorithm was also applied to the original data (without duplicates) to support the interpretability of some clusters generated by the t-SNE. Figure 6 shows how cluster 3 is mainly concentrated in the first quadrant of the third and fifth dimension (these were chosen because this is where the third cluster is most clearly represented). This indicates, according to the construction of each dimension of the PCA, that foods are concentrated to a lesser extent in areas where there are many trees, a lot of work and commerce, many street lights and schools, and few sidewalk ramps. Moreover, with regard to establishments dedicated to services, these represent approximately 10% in most of the clusters, except in cluster 3 where they represent 20.96%. Therefore, it can be said that establishments dedicated to services are more likely to be located in areas with characteristics that are not very advantageous for foods.

Finally, a main map with the results were displayed spatially. The map takes specifically the group of foods given that this was the largest of them all. In the same hexagonal grid as previous maps, it shows two zones, differentiated with degraded colors. The two zones had the most correlated categories of the PSAT indicator and the street vending foods (the variables with more correlation were jobs, housing, longitude of blocks, and schools). A green-colored zone named exploited conditions, contained both the correlated categories and ongoing food establishments. The second zone, purple colored, named potential conditions, had the correlated categories but no food establishments (SV-SF).

6. Discussion

This exercise analyzed the employment and economic phenomenon known as street vending, from the study case of the semi-fixed establishments in the Monterrey metropolitan area.
The exercise defined three general categories (food, services, and products), each with a particular ecosystem of urban characteristics. These constitute important factors in determining the success of new establishments, given the potential surge of the activity resulting from the SARS-CoV-2 factor in the world economy. Thus, the methodology can be replicated in cities with enough data and a considerable amount of the activity.

This result allows adequate policies in some ranges, the most important being the economy and employment, given the plausibility of permissions, sustained with a location strategy, and related topics such as urban zoning, housing, mobility, services, and others. This also opens the opportunity for private and social sectors to approach the phenomenon better, considering the street vending ecosystem characteristics to permit the activity’s success.

It also needs subsequent research and analysis, adding other variables as socioeconomic sectors, (in)security, and others. It should also consider adding the space–time geographical category, allowing analysis in tendencies and projections.

Software

There were two phases: first, the SV and PSAT process with Qgis. The information was processed with a grid generator and also a filter box to run Structured Query Language (SQL), creating the PSAT tool. For the second phase, we used, H2O, an open source for artificial intelligence. H2O runs Deep Learning, an artificial neural network that is trained. Furthermore, H2O runs gradient boosting, a technique for regression analysis and statistical classification problems, which produces a predictive model in the form of a set of weak prediction models, typically decision trees. Furthermore, we used R, a programming language and free software environment for statistical computing and graphics, which runs clustering, a grouping process to perform the phenomenon better, considering the street vending ecosystem characteristics to permit the activity’s success.

Acknowledgments

A. Barreda gratefully expresses gratitude to the Mexican National Council for Science and Technology (CONACYT) for the scholarship to pursue postgraduate studies. Furthermore, to Fernando Tovar Zamora Plowes and Carlos Abel Reséndiz Terrazas for the GIS expertise and assistance.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Antonio Alfonso Barreda Luna DOI https://orcid.org/0000-0002-2850-4644
Gonzalo Hatch Kuri DOI https://orcid.org/0000-0002-5189-4708
Juvenal Rodríguez-Reséndiz DOI https://orcid.org/0000-0001-8598-5600
Marco Antonio Zamora Antuñano DOI https://orcid.org/0000-0002-9865-3944
José Antonio Altamirano Corro DOI https://orcid.org/0000-0003-0061-091X

Data availability statement

The data that support the findings of this study are available from the corresponding author [JRR], upon reasonable request.

References

Batty, M. (2009). Accessibility: In search of a unified theory.Environment and Planning B: Planning and Design, 36(2), 191–194. https://doi.org/10.1068/b3602ed
Bjerkholt, O. (2016). Ragnar anton kittil frisch (1895–1973). In Handbook on the history of economic analysis (Vol. I).Edward Elgar Publishing.
Carpio-Pinedo, J., Benito-Moreno, M., & Lamíquiz-Daudén, P. J. (2021). Beyond land use mix, walkable trips. An approach based on parcel-level land use data and network analysis. Journal of Maps, 17(1), 23–30. https://doi.org/10.1080/17445647.2021.1873063
Contreras, V., & Welhert, U. (1988). Sobre vivir en la calle: El comercio ambulante en Santiago. PREALC.
Crossa, V. (2016). Editores invitados. Reading for difference on the street: De-homogenising street vending in Mexico City. Urban Studies, 53(2), 287–301. https://doi.org/10.1177/0042098014563471
Domínguez-Gijón, R. M., & Venegas-Martínez, F. (2016). Sector informal en méxico: Un modelo estocástico microeconómico del vendedor ambulante. Análisis Económico, 31(77), 27–52. http://analisiseconomico.azc.uam.mx/index.php/rae/article/view/50
Ehrenfeucht, R. (2017). Do food trucks and pedestrians conflict on urban streets? Journal of Urban Design, 22(2), 273–290. https://doi.org/10.1080/13574809.2017.1281731
García-Guzmán, J. (2001). Los vendedores ambulantes en la ciudad de méxico. Planteamiento para un modelo económático. Inédita. México. Tesis presentada para aspirar al grado de Licenciado en Economía. UNAM. Facultad de Economía. http://herzog.economia.unam.mx/seccs/docs/tesisfe/GarciaGJ/cap1-1.pdf
Glicker, J. J. (2014). Food cart economics: A comprehensive analysis of Portland’s street food market [PhD thesis]. University of Oregon.
HABITAT, U. (2019). Manual de calles. Diseño vial para Ciudades Mexicanas.http://onuhabitat.org.mx/index.php/manual-de-calles-diseno-vial-paraciudades-mexicanas.
Haggett, P., & Chorley, R. (1967). Modelos, paradigmas y la nueva geografía. In Modelos en geografía (pp. 19–41).London. Methuen and CO LTD.
Hardy, A. (1996). On the number of clusters. Computational Statistics & Data Analysis, 23(1), 83–96. https://doi.org/10.1016/S0167-9473(96)00022-9
Hart, K. (1973). Informal income opportunities and urban employment in Ghana. *The Journal of Modern African Studies, 11*(1), 61–89. https://doi.org/10.1017/S0022278X00008089

ILO, International Labour Organization (2020). Covid-19 crisis and the informal economy: Immediate responses and policy challenges. *ILO Brief.*

INEGI, Instituto Nacional de Estadística y Geografía (2010). *Censo de población y vivienda.* Technical report, INEGI.

INEGI, Instituto Nacional de Estadística y Geografía (2016,2018). *Inventario nacional de vivienda.* aguascalientes. Technical report, INEGI.

INEGI, Instituto Nacional de Estadística y Geografía (2019). *Resultados de la encuesta nacional de ocupación y empleo. aguascalientes.* Technical report, INEGI.

ITDP, Institute for Transportation and Development Policy (2018). Pedestrians first: A tool for walkable cities. New York: ITDP https://www.itdp.org/publication/walkability-tool/

Jacoby, E., & Pardo, C. F. (2012). Ciudades del automóvil, obesidad y cambio climático. Slow Research. http://www.bvsde.paho.org/teексcom/cd045364/ciudadesauto.pdf

John, C. R., Watson, D., Russ, D., Goldmann, K., Ehrenstein, M., Pitzalis, C., Lewis, M., & Barnes, M. (2020). M3C: Monte Carlo reference-based consensus clustering. *Scientific Reports,* 10(1), 1–14. https://doi.org/10.1038/s41598-019-56847-4

Krambeck, H. V. (2006). *The global walkability index* [PhD thesis]. Massachusetts Institute of Technology.

Lewis, J., & Draeger, M. (2011). *Food on wheels: The complete guide to starting a food truck, food cart, or other mobile food business.* Rabbit Ranch Publishing.

Linderman, G. C., & Steinerberger, S. (2019). Clustering with t-SNE, provably. *SIAM Journal on Mathematics of Data Science, 1*(2), 313–332. https://doi.org/10.1137/18M1216134

Meneses-Reyes, R. (2018). (Un)authorized: A study on the regulation of street vending in Latin America. *Law & Policy, 40*(3), 286–315. https://doi.org/10.1111/lapo.2018.40.issu e-3

NACTO, National Association of City Transportation Officials (2021). *Streets for Pandemic Response and Recovery.*

Newman, L. L., & Burnett, K. (2013). Street food and vibrant urban spaces: Lessons from Portland, Oregon. *Local Environment, 18*(2), 233–248. https://doi.org/10.1080/13549839.2012.729572

Ortega, E., Martín, B., De Isidro, Á., & Cuevas-Wizner, R. (2020). Street walking quality of the ‘Centro’ district, Madrid. *Journal of Maps, 16*(1), 184–194. https://doi.org/10.1080/17445647.2020.1829114

Pánek, J. (2019). Mapping citizens’ emotions: Participatory planning support system in Olomouc, Czech Republic. *Journal of Maps, 15*(1), 8–12. https://doi.org/10.1080/17445647.2018.1546624

Pirie, G. H. (1979). Measuring accessibility: A review and proposal. *Environment and Planning A: Economy and Space,* 11(3), 299–312. https://doi.org/10.1068/a110299

Pivo, G., & Fisher, J. D. (2011). The walkability premium in commercial real estate investments. *Real Estate Economics,* 39(2), 185–219. https://doi.org/10.1111/reec.2011.39.issue-2

Rodgers, K., & Roy, K. (2010). *Cartopia: Portland’s food cart revolution.* Roy Rodgers Press.

Rueda, S. et al. (2010). *Plan de indicadores de sostenibilidad urbana de vitoria-gasteiz.* Technical report. Agencia de Ecología Urbana de Barcelona, Ayuntamiento Vitoria Gasteiz.

Score, W. (2014). Walk score methodology. *Accessed April,* 24.

Silva Londoño, D. A. (2010). Comercio ambulante en el centro histórico de la ciudad de méxico (1990–2007). *Revista Mexicana de Sociología,* 72(2), 195–224. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0188-25032010000200001

Sofwan, M., & Tanjung, M. H. (2020). Evaluation study of street walking quality of the central business district (CBD) area, Pekanbaru city. *Journal of Geoscience, Engineering, Environment, and Technology,* 5(3), 175–185. https://doi.org/10.1068/a110299

Thatcher, O. J. (1907). The library of original sources (Vol. 3). University Research Extension.

Wattenberg, M., Viégas, F., & Johnson, I. (2016). How to use t-SNE effectively. *Distill, 1*(10), e2. https://doi.org/10.25299/distill