Validation of secondary data sources of the retail food environment in the capital of Uruguay, an emerging Latin American country

Leticia Vidal[1], Florencia Alcaire[1], Gerónimo Brunet[2], María Costa[3], Sofía Verdier[3], María Rosa Curutchet[3], Luciana Bonilla[3], Sergio Turra[4], Fernanda Risso[4], Leandro Machín,[5] Gastón Ares[1]

[1] Sensometrics & Consumer Science, Instituto Polo Tecnológico de Pando, Facultad de Química, Universidad de la República. By Pass de Rutas 8 y 101 s/n. CP 91000. Pando, Uruguay.

[2] Espacio Interdisciplinario, Universidad de la República, José Enrique Rodó 1843, CP 11200, Montevideo, Uruguay.

[3] Instituto Nacional de Alimentación, Ministerio de Desarrollo Social, Piedras 165, CP 11000, Montevideo, Uruguay.

[4] Escuela de Nutrición, Universidad de la República, Av. Ricaldoni S/N, CP 11600, Montevideo, Uruguay.

[5] Centro de Investigación Básica en Psicología, Facultad de Psicología, Universidad de la República. Tristán Narvaja 1674. CP 11200. Montevideo, Uruguay.

* Corresponding author: Leticia Vidal [lvidal@fq.edu.uy]

CRediT author statement

Leticia Vidal: Conceptualization, Methodology, Investigation (data collection), Analysis, Project administration, Funding acquisition, Writing – Review & Editing. Florencia Alcaire: Conceptualization, Methodology, Investigation (data collection), Writing – Review & Editing. Gerónimo Brunet: Conceptualization, Methodology, Investigation (data collection), Analysis, Writing – Review & Editing. María Costa: Conceptualization, Methodology, Investigation (data collection), Writing – Review & Editing. Sofía Verdier: Conceptualization, Methodology, Investigation (data collection), Writing – Review & Editing. María Rosa Curutchet: Conceptualization, Methodology, Investigation (data collection), Writing – Review & Editing. Luciana Bonilla: Conceptualization, Methodology, Writing – Review & Editing. Sergio Turra: Conceptualization, Methodology, Investigation (data collection), Writing – Review & Editing. Fernanda Risso: Conceptualization, Methodology, Investigation (data collection), Writing – Review & Editing. Leandro Machín: Investigation (data collection), Writing – Review & Editing. Gastón Ares: Conceptualization, Methodology, Analysis, Investigation (data collection), Writing – Original Draft, Writing – Review & Editing.

Funding

Financial support was obtained from Agencia Nacional de Investigación e Innovación (Project No. FCE_3_2022_1_172443), Instituto Nacional de Alimentación (Ministerio de Desarrollo Social, Uruguay), and Espacio Interdisciplinario (Universidad de la República, Uruguay).

Conflict of interest statement

The authors declare no conflicting interests.

Research highlights

- Field validation of secondary data sources of the retail food environment in Montevideo.
- 38% of the outlets found on the ground were not listed on any database.
- Sensitivity and concordance were moderate (0.606 and 0.488, respectively).
- Positive predictive value was substantial (0.702).
- Validity measures were positively associated with the socio-economic status of the census tract

Validation of secondary data sources of the retail food environment in the capital of Uruguay, an emerging Latin American country

Abstract

Secondary data sources are frequently used for characterizing physical access to food. Although several studies have shown that they tend to show a moderate agreement with field observation in WEIRD (Western Educated Industrialized Rich and Democratic) countries, little is known about their validity in non-WEIRD countries. The aim of the present research was to assess the validity of secondary data sources of the retail food environment in Montevideo, the capital of Uruguay, an emerging Latin American country. A random sample of 106 census tracts was obtained, covering 12% (62 km^2) of the city's total area. Two secondary data sources were considered: administrative records and Google maps. An aggregate database was created by manually removing duplicates. A total of 1051 unique outlets were listed in the database within the census tracts included in the sample. Field validation was performed by six teams of two observers. A total of 1217 open food outlets were identified on the ground, including 462 (38%) outlets not listed on any database. On the contrary, 296 outlets listed in the databases (28.2%) were not found or were closed at the time of field validation. At the aggregate level, sensitivity and concordance were moderate (0.606 and 0.488, respectively), whereas positive predictive value was substantial (0.702). However, large heterogeneity in the validity of the database across census tracts was found. Sensitivity, positive predictive value, and concordance were positively associated with the socio-economic status index of the census tract. These results suggest that secondary data sources must be used with caution, particularly for characterizing areas with low socio-economic status.

Keywords: *food environment; food access; food availability; food outlets; ground-truthing.*

1. Introduction

The food environment, defined as the physical, economic, political, and socio-cultural context through which consumers obtain, prepare, consume and discard food (HLPE, 2017), has a large influence on food choice, eating habits, and health outcomes (Konapur et al., 2022; Sawyer et al., 2021). Physical access to food is one of the dimensions of the food environment, which has received increasing attention from researchers and policy makers in the last decade (Turner et al., 2018). It refers to the location of food retail outlets and ease of reaching them, considering distance, travel time, and/or cost (Caspi et al., 2012; Penchansky & Thomas, 1981). Physical access to food has a key role for creating a supportive environment to promote healthy and sustainable eating patterns (Story et al., 2008).

Geographic analysis of food retail outlets based on secondary data sources, such as administrative records and commercial databases, has been the main methodological approach to characterize physical access to food (Lebel et al., 2017; Lytle & Sokol, 2017). This approach is often cost-effective, particularly when dealing with large geographic areas, and enables the analysis of historical data (Lytle & Sokol, 2017). However, secondary data sources have been shown to have important limitations in terms of validity and reliability (Forsyth et al., 2010; Lebel et al., 2017).

Although it is still not a common practice to assess the validity of secondary data sources in food environment research, several studies on the topic have been published (Lytle & Sokol, 2017). Reviews and meta-analyses of studies assessing the validity of commercially available business data has shown that the agreement with field observations, regarded as the gold standard, tends to be moderate (Fleischhacker et al., 2013; Lebel et al., 2017).

Most studies assessing the validity of secondary data sources have been developed in WEIRD countries (Canalia et al., 2020; Díez et al., 2019; Lebel et al., 2017; Lytle & Sokol, 2017). Yet, little is known about the validity of these data sources in non-WEIRD countries. Food retail environment research in these settings may face challenges related to the relevance of informal markets and to the lack of commercial data sources of the retail food environment (Turner et al., 2018). As far as it can be ascertained, only one study has been published assessing the validity of secondary data sources in an emerging country. A recent study has shown that the agreement of food retail data from Google Earth with field observations was moderate to excellent in two Brazilian cities: Rio de Janeiro and Belo Horizonte (de Menezes et al., 2021).

In this context, the aim of the present research was to assess the validity of secondary data sources of the retail food environment in Montevideo, the capital of Uruguay, an emerging Latin American country.

2. Methods

The study was conducted in Montevideo, the capital city of Uruguay, as part of a larger project aimed at characterizing the retail food environment. The project received approval from the ethics committee of (blinded for review) (Protocol No 101900-000043-22).

Montevideo has an estimated population of 1,670,545 inhabitants, living in an area of 526 km^2. Montevideo has an area of 526 km2 and a population of 1,670,545 inhabitants (Catálogo de datos geográficos de Montevideo, 2024). Approximately 321 km^2 (61% of the city's total area) is rural and 179 km^2 (34% of total area) is urban (Catálogo de datos geográficos de Montevideo, 2024). According to latest data, 12.8% of the city's population has a per capita income below the poverty line (Instituto Nacional de Estadística, 2024). The mean years of schooling for adults aged 25 or more is 11.4 (Instituto Nacional de Estadística, 2024). The city is administratively divided in 8 municipalities, which intend to deepen governmental management and democratic participation.

2.1. Sampling

Ten percent of the city's census tracts (N=1,063) were selected using probability proportional to size sampling, considering the number of census tracts in each of the eight municipalities. A sample of 106 census tracts was obtained. The selected tracts covered 62 km^2 (12% of the city's total area) and widely differed in geographical location and socio-economic status (Figure 1). The geographic and socio-economic characteristics of the census tracts are shown in the Supplementary material (Table 1).

2.2. Secondary data sources

Two secondary data sources were considered: administrative records and Google maps. According to Uruguayan legislation, all food outlets should be registered at national and/or local government, depending on the type of outlet. To capture all registered outlets, four different publicly available administrative records were considered: Register of Food Establishments from the Food Regulation Service of Montevideo local government, Single National Registry of Meat Establishments from the National Meat Institute, the National Register of Pharmacies from the Ministry of Public Health and register of farmers' markets from Montevideo local government. Table 1 shows an overview of the types of outlets retrieved from the different records. For the Register of Food Establishments from the Food Regulation Service of Montevideo local government, the classifications considered within each type of outlet are shown. The following information was retrieved from the administrative records: name of the outlet, address, type of outlet. Data were obtained in July-August 2023. Outlets were geolocated based on the address listed on the database, using the Unique System of Geographic Addresses of Uruguay on package sf (Pebesma, 2023) for R software (R Core Team, 2022). The great majority of the outlets (n=6490, 98%) were successfully geolocated using the automatic procedure. However,

125 outlets were manually geolocated using Google maps considering the address listed on the database, whereas 7 outlets could not be geolocated.

An automatic procedure was developed to retrieve information on the food outlets available in Montevideo using Google maps. The package googleway (Cooley & Barcelos, 2023) was used to access the Google maps API through R software (R Core Team, 2022). A grid of 272 points distributed along the city was created (Supplementary Material Figure 1). The resolution of the grid was adjusted to the area of the municipalities. For each point in the grid, searches were performed for different types of outlets, considering radius of 600 m. The keywords used in the searches to identify different types of outlets are shown in Table 1. The following information was retrieved for each food outlet: place_id (from the Google maps API), name of the outlet, address, geographical coordinates, and keyword used in the search. After removing outlets with duplicate place_id, a total of 7,086 outlets were identified.

For each of the databases (administrative records and Google maps), all the outlets located within the 106 census tracts included in the sample were identified using package sf (Pebesma, 2023) for R software (R Core Team, 2022). An aggregate database was created by manually removing the duplicates considering both the name and the address of the outlet.

2.3. Field validation

Six teams of two observers conducted field work to validate the secondary data sources. The observers received training on the validation instruments and on the use of maps of the census tracts. They were instructed to walk all the streets within the census tracts, observing both sides of the streets except for the boundaries of the tract. They were asked to register all the stores selling foods and beverages using a web form on Compusense Cloud (Compusense Inc., Guelph, Canada). Observers had a list of the outlets identified in the database for each census tract, identified by name, address and type of store. For the outlets listed on the database they were

asked to indicate the status (located and open, closed or not found, open but type of outlet changed, open but name changed), and to check the name of the outlet, address, and type of store.

For the outlets found on the ground but not listed on the database, they had to register the name, address and type of store. Observers entered all the outlets to gather information about the types of foods sold, number of cash registers for supermarkets and grocery stores, and where customers ordered for bars, restaurants and cafés. Field work was conducted on weekdays, from 9AM to 1 PM or from 3 PM to 5 PM, between September 2023 and February 2024. For farmers' markets, a part of the field work was conducted on weekend days, depending on their opening days.

2.4. Data analysis

The validity of the secondary data sources was assessed using three indicators: i) sensitivity: proportion of the open outlets observed on the ground and open that were listed in the database; ii) positive predictive value: proportion of the outlets listed in the database that were found on the ground and were open; iii) concordance: proportion of the outlets listed on the database that were found on the ground and were open among all the outlets (listed in the database and found on the ground but not listed). Food outlets that matched in address and type but with variations in name were regarded as correctly identified given that the difference introduced no changes in terms of food availability. On the contrary, outlets matching by address but corresponding to different types of stores were regarded as mismatched observations. For example, an outlet listed as a grocery store but functioning as a bar at the time of the field validation was regarded as a mismatch. Regarding address, outlets matching in name but differing in location in approximately 20 meters were regarded as a match.

The indicators were first calculated at the aggregate level. Then, they were calculated separately for each of the databases (administrative records and Googlemaps), the different types

of stores, and each of the census tracts. Stores were sorted in the following types: supermarekets and grocery stores; bars, restaurants and cafés; bakeries; kiosks or candy shops; fruit and vegetable stores or farmers' markets; pharmacies; butchers', poultry and fish stores; ice-cream stores; street vendors; other specialty stores (e.g., cheese stores; fresh pasta stores; delicatessen); other types of stores (e.g., nutritional supplements stores, herbalist's stores). The following criteria were used to interpret the quality of the indicators: 0.00-0.20 poor, 0.21-0.40 fair, 0.41-0.60 moderate, 0.61-0.80 substantial, 0.81-1.00 almost perfect (Landis & Koch, 1977; Lebel et al., 2017).

The existence of systematic bias related to the type of store and the socio-economic status of the census tracts was assessed. Linear regressions were run to explore associations between each of the three validity indicators and the socio-economic status index of the census tracts. A significance level of 5% was considered. Data analyses were performed using R software (R Core Team, 2022).

3. Results

As shown in Table 2, 730 food outlets were listed in the administrative records, whereas 655 outlets were identified in Google maps. A total of 334 outlets were duplicated, whereas 717 were identified in only one of the databases: 395 were identified only in the administrative records and 322 were identified only in Google maps. Bars, restaurants, and cafés were the most prevalent food outlet in the aggregate database (n=350, 33.3%), followed by supermarkets and grocery stores (n=263, 25.0%), and bakeries (n=117, 11.1%).

The field validation identified a total of 1217 open food outlets, including 361 supermarkets and grocery stores (29.6% of total open food outlets), 295 bars, restaurants and cafés (24.2%), 143 kiosks or candy shops (11.8%) and 108 bakeries (8.9%). As shown in Table 2, 462 food outlets were found on the ground but were not listed in any of the databases, accounting for 38.0%

of the total number of open food outlets. Supermarkets, grocery stores, and kiosks or candy shops were the types of outlets most frequently found on the ground but not listed in the databases (Table 2). In fact, 47.1% of the open supermarkets or grocery stores and 60.8% of the open kiosks or candy shops were not listed. Meanwhile, 296 outlets listed in the databases (28.2%) were not found or were closed at the time of the field validation. In addition, 1.6% of the food outlets included in the database were open but corresponded to a different type of outlet.

The validity statistics of the validation are shown in Table 2. The sensitivity of the aggregate database was 0.606. Thus, its ability to capture existing food outlets can be regarded as moderate. The positive predictive value of the aggregate database was 0.70, indicating a good verification rate, i.e., the outlets listed in the database had a good likelihood of being found open during the validation. Finally, the concordance of the aggregate database and the results of the field validation was also moderate (0.488).

Differences in the validity of the administrative records and Google maps were found. As shown in Table 2, the database created from administrative records had higher values of the three validity statistics than Google maps.

The validity measures of the aggregate database differed across types of food outlets. The largest sensitivity was found for pharmacies, and bars, restaurants and cafés. For these types of outlets, the values where higher than 0.8, suggesting almost perfect ability of the aggregate database to capture existing food outlets. On the contrary, the lowest sensitivity was found for street vendors, showing a poor ability to capture existing outlets (Table 2). For bakeries, and butchers', poultry and fish stores, the sensitivity can be regarded as substantial, whereas for the rest of the food outlets it corresponded to moderate. Concordance followed a similar pattern, whereas positive predictive power did not largely differ among the different types of food outlets with the exception of street vendors, other specialty stores, and other types of stores (Table 2). Similar results were found when differences were analyzed separately for the databases created from administrative records and Google maps (data not shown).

Large heterogeneity in the validity of the aggregate database across census tracts was found. The values of the three validity measures at the level of individual census tracts ranged between 0.000 and 1.000 (Figure 2). The interquartile range was 0.333-0.833 for sensitivity, 0.555-0.857 for positive predictive value, and 0.236-0.677 for concordance. The validity of the aggregate database tended to increase with the socio-economic status of the census tract (Figure 3). Significant linear associations were found between the three validity measures and the socio-economic status index of the census tract: sensitivity: $r=0.684$, $p<0.001$; positive predictive value: $r=0.502$, $p=0.002$; concordance: $r=0.647$, $p<0.001$.

4. Discussion

Secondary data sources are highly relevant for food environment research (Lebel et al., 2017; Lytle & Sokol, 2017). However, their validity has been mainly assessed in WEIRD countries. The present study aimed at contributing to fill a gap in retail food environment literature by assessing the validity of secondary data sources in Montevideo, the capital of Uruguay, a non-WEIRD Latin American country.

To the authors' knowledge this is the first study to assess the validity of secondary data sources based on administrative records in the Latin American context and one of the few to combine administrative records with the information available on Google maps. Although the results show that the sensitivity of administrative records can be regarded as moderate (Landis & Koch, 1977; Lebel et al., 2017), they also show that this database missed more than half the outlets found open on the ground. The estimates of sensitivity, positive predictive value and concordance of the administrative records were in the lower range of those reported by studies conducted in WEIRD countries using administrative records or commercial databases (Canalia et al., 2020; Daepp & Black, 2017; Díez et al., 2019; Lebel et al., 2017; Lucan et al., 2020). The relative lower sensitivity compared to studies conducted in WEIRD countries can be

explained by the existence of informal stores, which have been shown to be relevant in the Latin American context (Farah et al., 2023).

Google maps had a slightly lower validity, which can be categorized as fair. The sensitivity, positive predictive value and concordance for this database were markedly lower than those reported by (de Menezes et al., 2021) for two Brazilian cities. Although Google maps and other open geocoding services are being increasingly used in public health and epidemiological research to characterize features of the built environment (Lemke et al., 2015; Präger et al., 2019; Silva et al., 2015), results from the present work stress the need to assess their validity on a case by case basis to avoid misleading conclusions.

The combination of the two data sources led to an improvement in validity, increasing both sensitivity and concordance. Similar results have been reported by (Liese et al., 2010) when combining different commercial databases in the USA. The validity, positive predictive value, and concordance of the aggregate database was moderate, as reported by studies conducted in WEIRD countries using administrative records or commercial databases (Canalia et al., 2020; Daepp & Black, 2017; Díez et al., 2019; Lebel et al., 2017). These results suggest that Google maps seems to be a cost-effective approach to complement administrative records or commercial databases. Its main disadvantage is that historical data cannot be obtained.

Systematic differences in the validity of the secondary data sources across types of outlets were found. This result is highly relevant for food environment research given that different types of outlets imply differences in the relative availability of healthy and unhealthy food (Charreire et al., 2010; Gebremariam et al., 2017; Titis et al., 2022; Ver Ploeg et al., 2015). In the present work, sensitivity and concordance for supermarkets and grocery stores, as well as fruit and vegetable stores were lower than the aggregate values for all food outlets. Therefore, the use of secondary data sources may lead to measurement error by underestimating the availability of

healthy foods, such as fruits and vegetables. This result has not been found in previous studies although reports of differences in the validity of secondary data sources among different types of outlets exist. (Powell et al., 2011) reported higher sensitivity for supermarkets and grocery stores compared to other types of outlets. More recently, (Lucan et al., 2020) reported that government records underestimate the presence of street vendors and other categories of food outlets, such as barbers, laundromats, and newsstands, which were sources of both healthy and unhealthy foods.

Results from the present work also showed that the validity of the secondary data sources was largely heterogeneous across census tracts. A systematic variation with the socio-economic status of the areas was found. Sensitivity, positive predictive value, and concordance tended to significantly increase with the socio-economic status index of the census tract; i.e., secondary data sources were less able to accurately capture existing open food outlets in low socio-economic status areas compared to areas with high socio-economic status. This result may be explained by a larger proportion of informal food outlets in areas of low socio-economic status, which are not included in the administrative records. Therefore, secondary data sources to characterize the retail food environment may lead to inaccurate results in vulnerable areas of the city, which are more likely to be affected by inadequate physical access to food (Ver Ploeg et al., 2012).

Results from the present work are not in agreement with previous studies conducted in WEIRD countries, reporting no consistent associations between the validity of secondary data sources and socio-economic characteristics of the areas (Daepp & Black, 2017; Díez et al., 2019; Lebel et al., 2017; Liese et al., 2010; Powell et al., 2011). Results are also not in agreement with the only study assessing the validity of secondary data sources in a non-WEIRD country. The study was conducted in two Brazilian cities also did not find evidence of systematic differences in the validity of Google Earth data with socio-economic status (de Menezes et al.,

2021). Lack of consistence between studies conducted in the same region stresses the need for additional research.

Finally, it is worth mentioning that, despite the fact that census tracts were selected using random sampling, field work was restricted to 12% of the city's area, hindering the precision and accuracy of the estimates of validity indicators. This limitation applies not only to the observed differences among the values of validity indicators across types of food outlets, but also to generalizations of results to the total area of the city.

5. Conclusions

The validity of secondary data sources to characterize the sampled areas of the retail food environment of Montevideo was moderate. Systematic differences in the validity of the secondary data sources with socio-economic status of the census tracts were found. Based on the results, the use of secondary data sources does not seem a valid approach to characterize the retail food environment of areas with low socio-economic status in Montevideo. The dynamic nature of the food environment and the relevance of informal outlets in non-WEIRD countries suggest that these results may be extrapolated to similar settings. However, additional research is needed to confirm this hypothesis.

References

Canalia, C., Pinho, M., Lakerveld, J., & Mackenbach, J. (2020). Field Validation of Commercially Available Food Retailer Data in the Netherlands. *International Journal of Environmental Research and Public Health*, *17*(6), 1946. https://doi.org/10.3390/ijerph17061946

Caspi, C. E., Sorensen, G., Subramanian, S. V., & Kawachi, I. (2012). The local food environment and diet: A systematic review. *Health and Place*, *18*(5), 1172–1187. https://doi.org/10.1016/j.healthplace.2012.05.006

Catálogo de datos geográficos de Montevideo. (2024, April 17). *Zonificación primaria ultima del POT (Antecedente histórico)*. Https://Geoweb.Montevideo.Gub.Uy/Geonetwork/Srv/Spa/Catalog.Search#/Metadata/67495643-4ad2-4894-Ba0d-4b5bebe769c8.

Charreire, H., Casey, R., Salze, P., Simon, C., Chaix, B., Banos, A., Badariotti, D., Weber, C., & Oppert, J.-M. (2010). Measuring the food environment using geographical information systems: a methodological review. *Public Health Nutrition*, *13*(11), 1773–1785. https://doi.org/10.1017/S1368980010000753

Cooley, D., & Barcelos, P. (2023). *Package 'googleway'*. CRAN.

Daepp, M. I., & Black, J. (2017). Assessing the validity of commercial and municipal food environment data sets in Vancouver, Canada. *Public Health Nutrition*, *20*(15), 2649–2659. https://doi.org/10.1017/S1368980017001744

de Menezes, M. C., de Matos, V. P., de Pina, M., Costa, B. V. L., Mendes, L. L., Pessoa, M. C., Souza-Junior, P. R. B., Friche, A. A. L., Caiaffa, W. T., & Cardoso, L. O. (2021). Web Data Mining: Validity of Data from Google Earth for Food Retail Evaluation. *Journal of Urban Health*, *98*, 285–295. https://doi.org/10.1007/s11524-020-00495-x

Díez, J., Cebrecos, A., Galán, I., Pérez-Freixo, H., Franco, M., & Bilal, U. (2019). Assessing the Retail Food Environment in Madrid: An Evaluation of Administrative Data against Ground Truthing. *International Journal of Environmental Research and Public Health*, *16*(19), 3538. https://doi.org/10.3390/ijerph16193538

Farah, I., Stern, D., Ramírez, Y., López-Olmedo, N., Pérez-Ferrer, C., Langellier, B. A., Colchero, M. A., & Barrientos-Gutierrez, T. (2023). Food and beverage purchases at formal and informal outlets in Mexico. *Public Health Nutrition*, *26*(5), 1034–1043. https://doi.org/10.1017/S1368980022002324

Fleischhacker, S. E., Evenson, K. R., Sharkey, J., Pitts, S. B. J., & Rodriguez, D. A. (2013). Validity of secondary retail food outlet data: a systematic review. *American Journal of Preventive Medicine*, *45*(4), 462–473. https://doi.org/10.1016/j.amepre.2013.06.009

Forsyth, A., Lytle, L., & Van Riper, D. (2010). Finding food: Issues and challenges in using Geographic Information Systems (GIS) to measure food access. *Journal of Transport and Land Use*, *3*(1). https://doi.org/10.5198/jtlu.v3i1.105

Gebremariam, M. K., Vaqué-Crusellas, C., Andersen, L. F., Stok, F. M., Stelmach-Mardas, M., Brug, J., & Lien, N. (2017). Measurement of availability and accessibility of food among youth: a systematic review of methodological studies. *International Journal of Behavioral Nutrition and Physical Activity*, *14*(1), 22. https://doi.org/10.1186/s12966-017-0477-z

HLPE. (2017). *Nutrition and food systems. A report by the High Level Panel of Experts on Food Security and Nutrition of the Committee on World Food Security*. Committee on World Food Security.

Instituto Nacional de Estadística. (2024, April 17). *Encuesta Continua de Hogares, año 2023*. Https://Www5.Ine.Gub.Uy/Documents/Demograf%C3%ADayEESS/HTML/ECH/Pobreza/2023/Estimaci%C3%B3n%20de%20la%20pobreza%20por%20el%20m%C3%A9todo%20del%20ingreso%20anual%202023.Html.

Konapur, A., Gavaravarapu, S. R. M., & Nair, K. M. (2022). The 5 A's Approach for Contextual Assessment of Food Environment. *Journal of Nutrition Education and Behavior*, *54*(7), 621–635. https://doi.org/10.1016/j.jneb.2022.02.017

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Lebel, A., Daepp, M. I. G., Block, J. P., Walker, R., Lalonde, B., Kestens, Y., & Subramanian, S. V. (2017). Quantifying the foodscape: A systematic review and meta-analysis of the validity of commercially available business data. *PLOS ONE*, *12*(3), e0174417. https://doi.org/10.1371/journal.pone.0174417

Lemke, D., Mattauch, V., Heidinger, O., & Hense, H. (2015). Wer trifft ins Schwarze? Ein qualitativer Vergleich der kostenfreien Geokodierungsdienste von Google und OpenStreetMap. *Das Gesundheitswesen*, *77*(08/09), e160–e165. https://doi.org/10.1055/s-0035-1549939

Liese, A. D., Colabianchi, N., Lamichhane, A. P., Barnes, T. L., Hibbert, J. D., Porter, D. E., Nichols, M. D., & Lawson, A. B. (2010). Validation of 3 Food Outlet Databases: Completeness and Geospatial Accuracy in Rural and Urban Food Environments. *American Journal of Epidemiology*, *172*(11), 1324–1333. https://doi.org/10.1093/aje/kwq292

Lucan, S. C., Maroko, A. R., Abrams, C., Rodriguez, N., Patel, A. N., Gjonbalaj, I., Schechter, C. B., & Elbel, B. (2020). Government data v . ground observation for food-environment assessment: businesses missed and misreported by city and state inspection records. *Public Health Nutrition*, *23*(8), 1414–1427. https://doi.org/10.1017/S1368980019002982

Lytle, L. A., & Sokol, R. L. (2017). Measures of the food environment: A systematic review of the field, 2007–2015. *Health & Place*, *44*, 18–34. https://doi.org/10.1016/j.healthplace.2016.12.007

Pebesma, E. (2023). *Package 'sf'*. CRAN.

Penchansky, R., & Thomas, J. W. (1981). The concept of access: Definition and relationship to consumer satisfaction. *Medical Care*, *19*(2), 127–140. https://doi.org/10.1097/00005650-198102000-00001

Powell, L. M., Han, E., Zenk, S. N., Khan, T., Quinn, C. M., Gibbs, K. P., Pugach, O., Barker, D. C., Resnick, E. A., Myllyluoma, J., & Chaloupka, F. J. (2011). Field validation of secondary commercial data sources on the retail food outlet environment in the U.S. *Health & Place*, *17*(5), 1122–1131. https://doi.org/10.1016/j.healthplace.2011.05.010

Präger, M., Kurz, C., Böhm, J., Laxy, M., & Maier, W. (2019). Using data from online geocoding services for the assessment of environmental obesogenic factors: a feasibility study. *International Journal of Health Geographics*, *18*(1), 13. https://doi.org/10.1186/s12942-019-0177-9

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Sawyer, A. D. M., van Lenthe, F., Kamphuis, C. B. M., Terragni, L., Roos, G., Poelman, M. P., Nicolaou, M., Waterlander, W., Djojosoeparto, S. K., Scheidmeir, M., Neumann-Podczaska, A., & Stronks, K. (2021). Dynamics of the complex food environment underlying dietary intake in low-income groups: a systems map of associations extracted from a systematic umbrella literature review. *International Journal of Behavioral Nutrition and Physical Activity*, *18*(1), 96. https://doi.org/10.1186/s12966-021-01164-1

Silva, V., Grande, A., Rech, C., & Peccin, M. (2015). Geoprocessing via Google Maps for Assessing Obesogenic Built Environments Related to Physical Activity and Chronic Noncommunicable Diseases: Validity and Reliability. *Journal of Healthcare Engineering*, *6*(1), 41–54. https://doi.org/10.1260/2040-2295.6.1.41

Story, M., Kaphingst, K. M., Robinson-O'Brien, R., & Glanz, K. (2008). Creating healthy food and eating environments: Policy and environmental approaches. *Annual Review of Public Health*, *29*, 253–272. https://doi.org/10.1146/annurev.publhealth.29.020907.090926

Titis, E., Procter, R., & Walasek, L. (2022). Assessing physical access to healthy food across United Kingdom: A systematic review of measures and findings. *Obesity Science and Practice*, *8*(2), 233–246. https://doi.org/10.1002/osp4.563

Turner, C., Aggarwal, A., Walls, H., Herforth, A., Drewnowski, A., Coates, J., Kalamatianou, S., & Kadiyala, S. (2018). Concepts and critical perspectives for food environment research: A global framework with implications for action in low- and middle-income countries. *Global Food Security*, *18*, 93–101. https://doi.org/10.1016/j.gfs.2018.08.003

Ver Ploeg, M., Breneman, V., Dutko, P., Williams, R., Snyder, S., Dicken, C., & Kaufman, K. (2012). *Access to Affordable and Nutritious Food: Updated Estimates of Distance to Supermarkets Using 2010 Data.* U.S. Department of Agriculture.

Ver Ploeg, M., Dutko, P., & Breneman, V. (2015). Measuring food access and food deserts for policy purposes. *Applied Economic Perspectives and Policy*, *37*(2), 205–225. https://doi.org/10.1093/aepp/ppu035

406 **Tables**

407

408 **Table 1.** Types of outlets retrieved from the different databases. For those including a diversity

409 of outlets, the keywords (in Spanish) used to identify each type of outlet are shown.

Database	Type of outlet	Keyword in Spanish
Register of Food Establishments from the Food Regulation Service of Montevideo local government	Supermarkets and grocery stores	almacén, autoservice, supermercado, venta de alimentos libres de gluten, venta de productos envasados
	Bars, restaurants and cafés	bar, cafetería, cantina, elaboración y venta de cerveza artesanal, carro de venta de chorizos y hamburguesas, centro comercial gastronómico, mercado gastronómico, salón de té, cocción y venta de empanadas, cocción y venta de pizzas, comidas al paso, elaboración de creps, elaboración y venta de churros, elaboración y venta de empanadas, elaboración y venta de hamburguesas, elaboración y venta de panchos, elaboración y venta de sushi, minutas, parrillada, pizzería, restaurante, servicio de buffet
	Bakeries	confitería, panadería, rotisería, sucursal de panadería, venta de donas, venta de postres, venta de productos de confitería, venta de productos de panadería, venta de productos de repostería, venta de productos de rotisería
	Kiosks or candy shops	expendio de maíz acaramelado y salado, máquina expendedora de chocolate, venta de golosinas, elaboración y venta de maíz acaramelado
	Fruit and vegetable stores	Venta de frutas y hortalizas
	Poultry and fish stores	pescadería, venta de productos elaborados a base de pollo, venta de pescado y mariscos
	Ice-cream stores	expendio de helado, máquina expendedora de helados, venta de helados envasados, heladería
	Other specialty stores (e.g., cheese stores; fresh pasta stores; delicatessen)	carro de venta de chacinados, fiambrería, licorería, sucursal de fábrica de pastas, venta de bebidas alcohólicas, venta de bebidas analcohólicas, venta de chorizos, venta de frutos secos, venta de hielo, venta de huevos, venta de lácteos y derivados, venta de pastas, venta de productos chacinados, venta de productos congelados, venta de productos de granja, venta de quesos, pastas frescas, elaboración y venta de chorizos,
Single National Registry of Meat Establishments from the National Meat Institute	Butcher's and poultry stores	N/A
National Register of Pharmacies from the Ministry of Public Health	Pharmacies	N/A
Register of farmers' markets from Montevideo local government	Farmers' markets	N/A

Google maps	Supermarkets and grocery stores	supermercado, autoservicio, almacén, tienda de alimentación
	Bars, restaurants and cafés	bar, restaurante, cafetería, pizzería, cervecería, parrilla, carrito, cantina, comida rápida
	Bakeries	panadería, confitería, respostería, rotisería
	Kiosks or candy shops	kiosko, golosinas, chocolate
	Fruit and vegetable stores or farmers' markets	frutas y verduras, frutería
	Butchers', poultry and fish stores	carnicería, pollería, pescadería
	Pharmacies	farmacias
	Ice-cream stores	heladería
	Street vendors	tortas fritas
	Other specialty stores (e.g., cheese stores; fresh pasta stores; delicatessen)	fiambrería, licorería, pastas, quesería, frutos secos, congelados, tienda de vinos

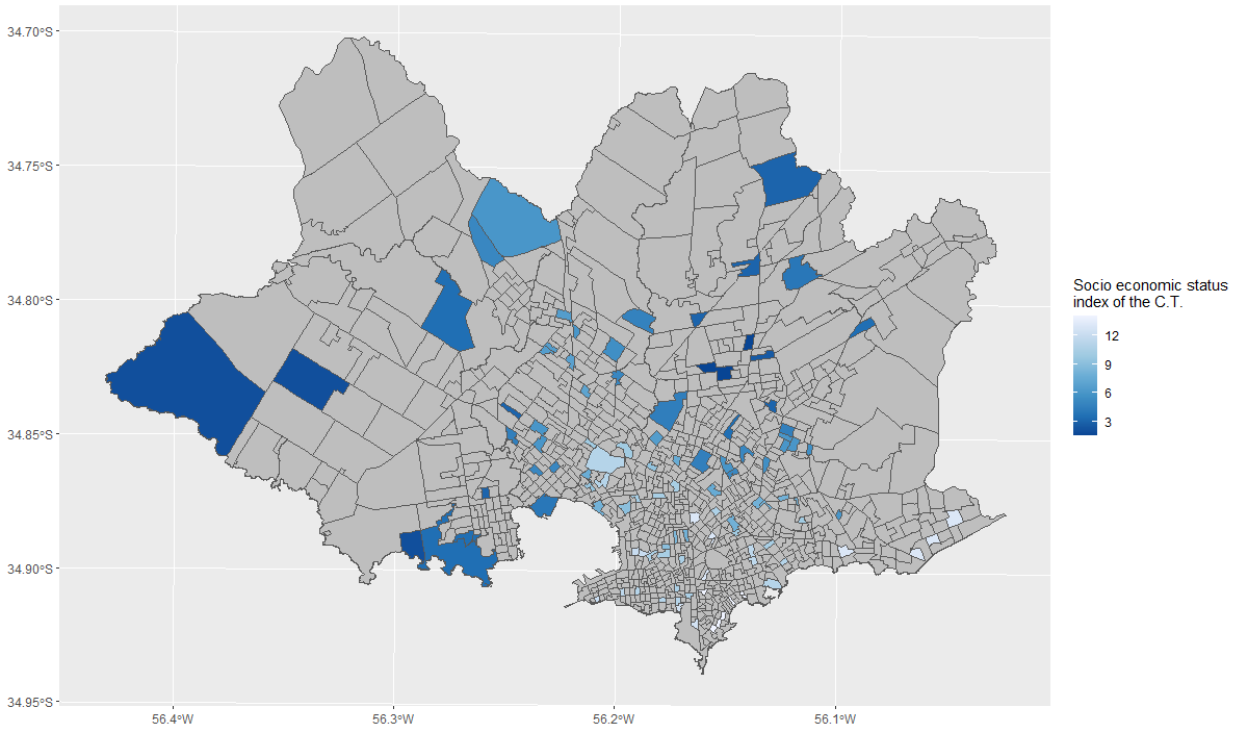
411 **Table 2.** Number of outlets listed on secondary data sources and results of the field validation conducted in 106 census tracts in

412 Montevideo (Uruguay), per database and type of outlet.

	Number of outlets listed in the database	Distribution of the outlets of the database according to the field validation (%)			Number of outlets found but not listed	Number of outlets found on the ground	Sensitivity	Positive predictive value	Concordance
		Located	Closed or not found	Located but type of outlet changed					
Database									
Administrative records	730	80.5	17.9	1.5	618	588	0.483	0.805	0.389
Google maps	655	71.3	27.3	1.4	741	467	0.384	0.713	0.309
Aggregate	1051	70.2	28.2	1.6	462	1217	0.606	0.702	0.488
Type of outlet									
Bars, restaurants and cafés	350	69.4	29.7	0.9	55	295	0.824	0.694	0.600
Supermarkets and grocery stores	263	72.6	26.2	1.1	170	361	0.529	0.726	0.441
Bakeries	117	71.8	27.4	0.9	24	108	0.778	0.718	0.596
Kiosks or candy shops	85	68.2	29.4	2.4	87	143	0.406	0.682	0.337
Fruit and vegetable stores or farmers' markets	59	67.8	25.4	6.8	49	89	0.449	0.678	0.370
Pharmacies	54	85.2	14.8	0	0	46	1.000	0.852	0.852
Butchers', poultry or fish stores	40	70	30	0	12	38	0.737	0.700	0.538
Other specialty stores	32	59.4	40.6	0	13	32	0.594	0.594	0.422
Other types of stores	31	51.6	35.5	12.9	13	29	0.552	0.516	0.364
Ice-cream stores	13	76.9	23.1	0	11	21	0.476	0.769	0.417
Street vendors	6	33.3	66.7	0	29	31	0.065	0.333	0.057

413

Figure 1. Socio-economic status index score for the census tracts (C.T.) included in the sample (n=106).



Note: Census tracts colored with grey were not included in the sample.

Figure 2. Sensitivity, positive predictive power and concordance of the aggregate database of the 106 census tracts in Montevideo (Uruguay)

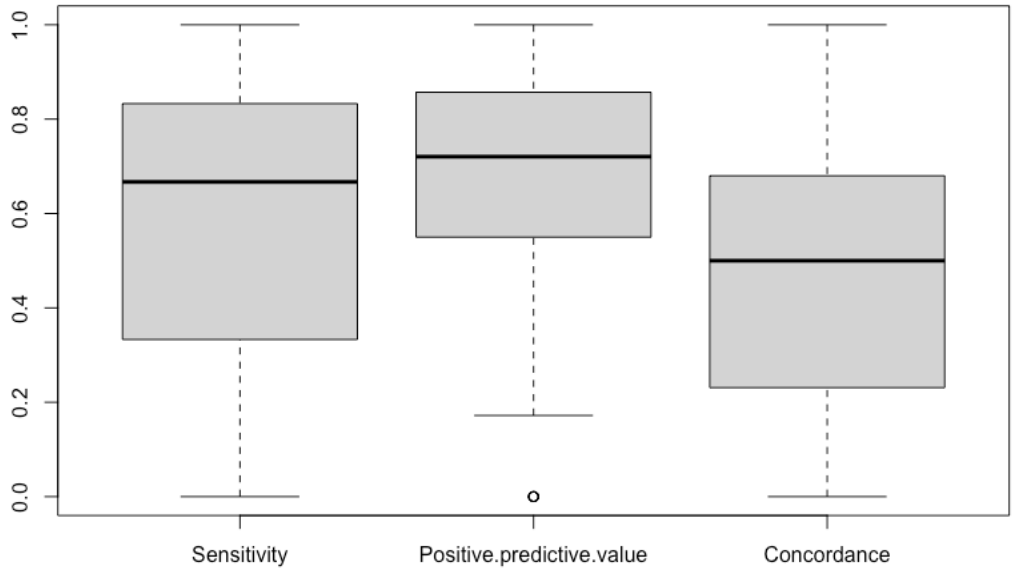
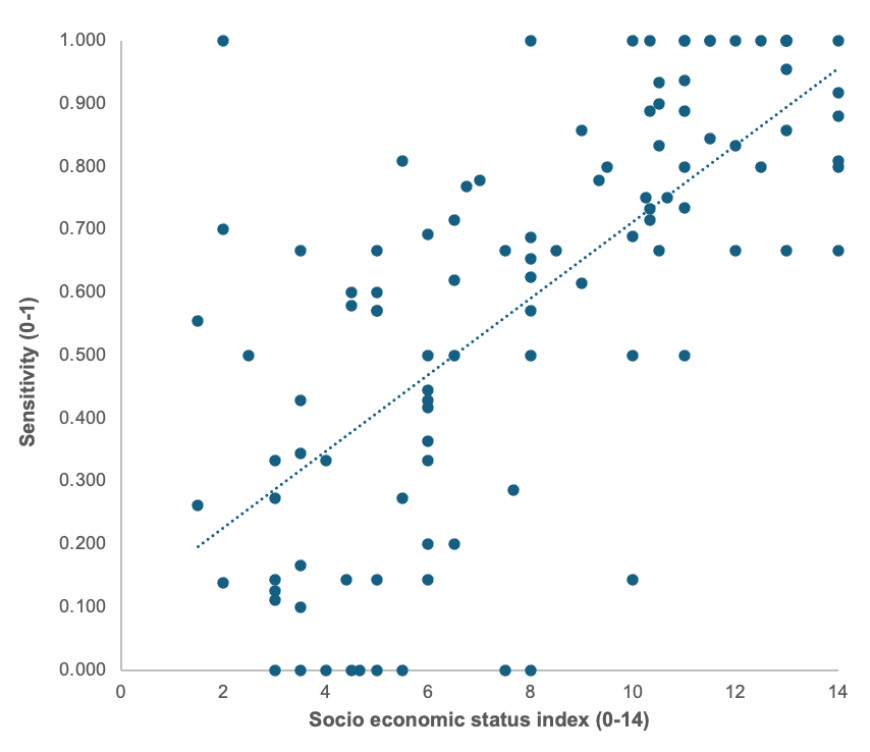
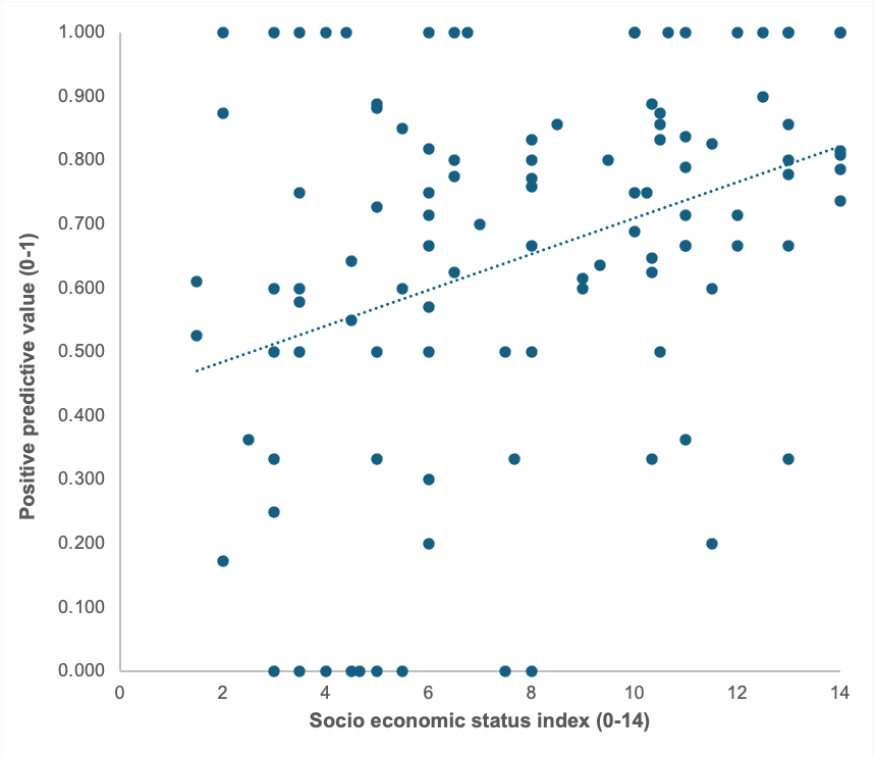


Figure 3. Sensitivity **(a)**, positive predictive value**(b),** and concordance **(c)** of the aggregate database as a function of the socio economic status index of the 106 census tracts in Montevideo (Uruguay).

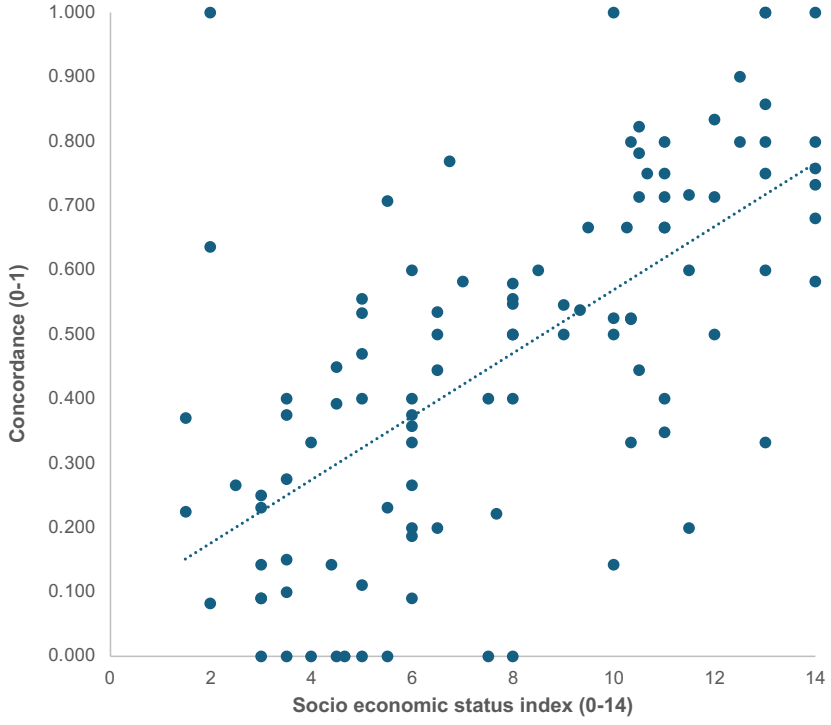
(a)



(b)


Positive predictive value (0-1)
1.000
0.900
0.800
0.700
0.600
0.500
0.400
0.300
0.200
0.100
0.000
0
2
4
6
8
10
12
14
Socio economic status index (0-14)

(c)


Concordance (0-1)
1.000
0.900
0.800
0.700
0.600
0.500
0.400
0.300
0.200
0.100
0.000
0
2
4
6
8
10
12
14
Socio economic status index (0-14)

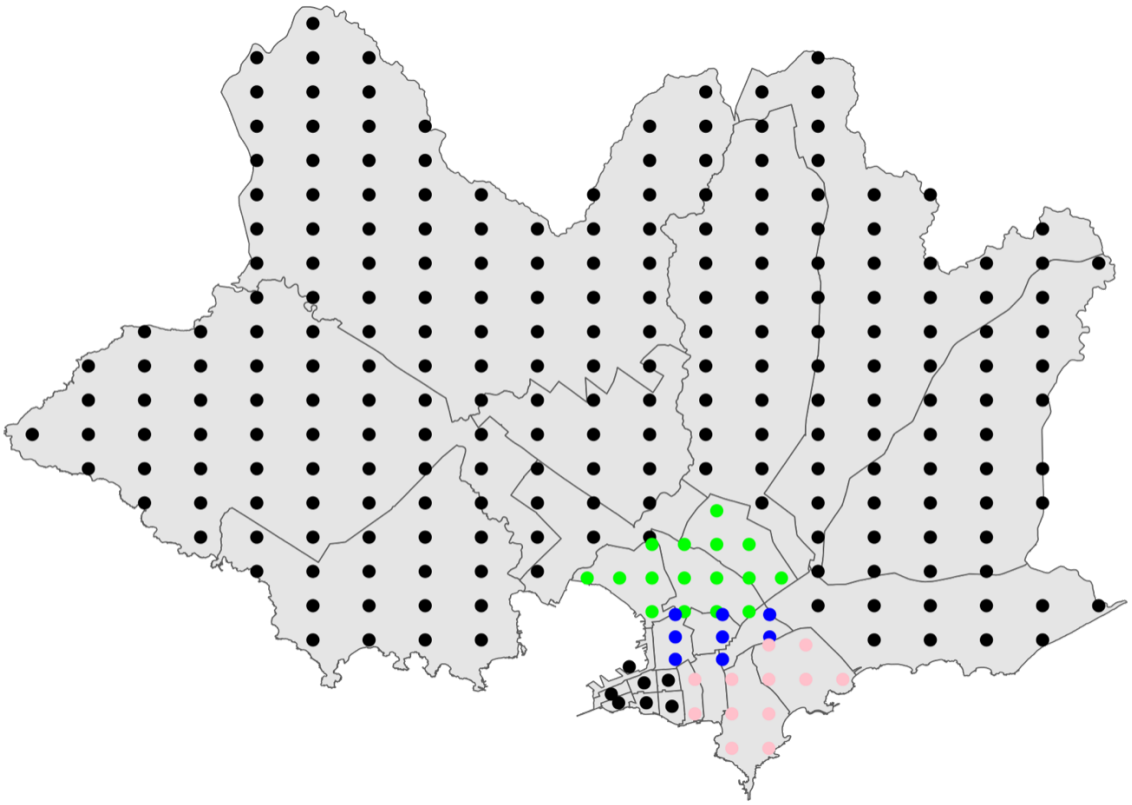
Supplementary Material

Table 1. Socio-economic characteristics of the census tracts of the sample (n=106).

Characteristic	Minimum	Mean	Maximum
Total area of the census tract (km^2)	0.0	0.6	18.8
Percentage of rural area in the census tract (%)	0.0	9.3	100.0
Total population in private households of the census tract	0.0	1220.3	5719.0
SES index of the census tract*	1.5	7.9	14.0
Per capita household income (USD)**	604.7	1095.2	2035.5
Households under the poverty line (%)**	0.0	8.5	24.5
Mean years of schooling of adults of the household**	8.2	11.4	15.1

* SES indexes of the census tracts were imputed from values of SES indexes for neighborhoods. In cases where one census tract was included in two or more neighborhoods, a simple average of the indexes was calculated. ** Statistics calculated for larger census areas (census sections), where census tracts are located. 1 USD = 39 Uruguayan pesos.

Figure 1. Grid of points used to perform the search of food outlets in Google Maps.



Note: Black countours correspond to the boundaries of the municipaplities. Different colours indicate different density of points according to the area of the municipalities.