Cureus Journal of Business and Economics

Open Access Original Article

The Effect of Store Location Selection Criteria on Sales Forecast in Brazil

Vanessa M. Santos [1] , Adrian K. Cernev [1] , Leonardo F. Ferraz [2] , Carlos E. Lourenco [2]

1. Technology and Data Science, Fundacao Getulio Vargas, Sao Paulo, BRA 2. Marketing, Fundacao Getulio Vargas, Sao Paulo, BRA

Corresponding author: Carlos E. Lourenco, caerib@gmail.com

Received 04/22/2025
Review began 04/24/2025
Review ended 05/21/2025
Published 06/17/2025

DOI: https://doi.org/10.7759/s44404-025-04990-9

Abstract

Many factors must be considered when an entrepreneur starts a business. Among them, a location that involves long-term commitment and high investments, especially for small and medium businesses, is of great importance. To examine this further, this paper measured the effect of store location selection criteria as applied to a chain of small stores situated on shopping streets in Brazil. This study examines the location importance model for retailers. By utilizing robust empirical data from a unique case study, this research extends our understanding of location selection criteria. Simple regression analysis is employed to assess the individual impact of each variable, while multiple regression analysis is used to predict store performance for both existing and new locations. The major finding of this paper is the proposition of a reasoned method for small and medium-sized retailers to evaluate the effects of influential factors in their decision-making process for choosing a store location. The research assessed the impact of various selection criteria on store performance forecasts. This research contributes to the literature by expanding knowledge about location selection criteria through the perspective of small business entrepreneurs, and it presents a clear and systematic method tailored for small and medium-sized retailers. This study validates a predictive model that can be replicated in the organizational contexts of other comparison goods retailers.

Categories: Consumer Behavior, Marketing and Market Issues in Business
Keywords: emerging markets, retail, sales forecast, small businesses, store location

Introduction

Entrepreneurs operating small businesses with limited resources and investments must remain aware of evolving consumer buying behaviors and trends. In recent years, a noteworthy transformation has been the substantial growth of e-commerce. This surge can be attributed to several factors, with particular emphasis on the widespread adoption of technology and the profound impact of the COVID-19 pandemic (NIELSENIQ, 2022). E-commerce is a highly convenient avenue for customers to shop, while for merchants, it presents a strategic means to attract a broader customer base. Nevertheless, it is important to underline that traditional offline commerce still retains its significance. Physical stores play an essential role in shaping an effective online sales strategy, fostering customer loyalty, and driving recurring sales in the realm of e-commerce (Economist, 2021).

Despite the steady growth of e-commerce, physical retail continues to play a central role in global commerce. Recent studies estimate that approximately 70% of global retail sales still occur in physical stores, underscoring their resilience and enduring relevance in consumer experience (Alexander and Blázquez-Cano, 2025). Moreover, the integration of digital technologies-such as artificial intelligence and augmented reality-has transformed brick-and-mortar environments into increasingly interactive and personalized spaces, strengthening the relationship between brands and consumers (Alaali, 2025).

Businesses that were exclusively online started to invest in the offline segment by opening concept stores and prioritizing a more human shopping experience (Zhang et al., 2019). However, the efficacy of physical stores hinges on their strategic positioning, which directly influences their ability to attract demand (Adeniyi et al., 2020) and ensure a convenient shopping experience (Chen and Tsai, 2016), especially for small retailers (Hensel et al., 2021) (Litz and Rajaguru, 2008). The process of selecting a prime location for a physical store involves various criteria and demands a comprehensive understanding of the factors affecting store performance (Mendes and Themido, 2003) (Turhan et al., 2013). Multiple analytical techniques and models have been devised to evaluate these aspects (Kang, 2016) (Shaikh et al., 2021). Despite the existence of more sophisticated methods (Chen and Tsai, 2016) (Ho et al., 2013), their adoption within the Brazilian retail landscape, especially by small and medium-sized enterprises (SMEs), remains limited (Deliberali et al., 2019) (Parente et al., 2012).

In smaller companies, the adopted techniques are more straightforward, such as the checklist and the analogy with similar stores. Advanced predictive tools, such as geographical information systems, are predominantly reserved for larger retail chains. However, the decision-making process must be grounded.

Choices about location involve long-term commitment and high investments, which can pose a financial risk, especially to small and medium-sized companies. It is essential to consider that the performance of small retailers is notably elevated when they operate in densely populated areas with easy accessibility (Litz and Rajaguru, 2008) (Seong et al., 2021). The objective of this paper is to comprehensively understand the location selection criteria that impact sales forecasts and how entrepreneurs can use these criteria in decision-making. Additionally, it aims to shed light on how entrepreneurs can effectively employ these criteria in their decision-making processes. Thus, this study addresses the following research question, RQ1: How do store location selection criteria affect new store sales forecasts?

For this study, a popular Brazilian and affordable clothing retailer was selected. This retailer brings the main characteristics desired for this study, such as size, several street stores, and target audience. This study investigated the effect of the store location selection criteria applied to a chain of small stores situated on shopping streets, using simple regression to analyze the impact of each variable and multiple regression to predict store performance for existing and new ones and provide criteria for managers to make decisions based on a straightforward and reasoned method. This Brazilian retail chain was selected due to its relevance as a representative SME with a strong presence in street commerce, an active expansion plan, and accessible, reliable data for modeling sales forecasts based on location criteria.

This research contributes to the literature in three ways. First, it expands knowledge about location selection criteria through the perspective of small business entrepreneurs, supported by unique and robust data collection. Second, it validates a predictive model that can be replicated in the organizational contexts of other comparison goods retailers. Third, it offers a straightforward and reasoned method for small and medium-sized retailers.

Store location selection criteria

Store location selection criteria are utilized to understand all aspects of physical store performance. Some of the challenges of offline retail in relation to location choice include fierce competition, economic uncertainties, online and multichannel retail, long-term commitment, and high investments, making it necessary for the decision-making process to be grounded. To mitigate these risks, companies must ascertain the most suitable model for their specific retail context, considering factors like speed, accuracy, and the resources at their disposal. It is essential to identify and evaluate the key determinants that impact the store's performance with discernment. Mendes and Themido suggest a classification of internal factors (Mendes and Themido, 2003), such as store size, store configuration and location variables, and external factors such as demand and competition. Turhan and colleagues (Turhan et al., 2013) reviewed the literature and formalized a theoretical model to categorize the main store location selection criteria: performance measures, population structure, economic factors, competition, saturation level, store characteristics, and magnet. A discussion of the theoretical framework and hypotheses are presented below.

Influence Area

Influence areas refer to the territorial extensions where the customers of a retailer or a retail cluster are predominantly located (Dolega et al., 2016). These influence areas are categorized based on their significance: primary (constituting approximately 60% to 75% of customers), secondary (comprising 15% to 25% of customers), and tertiary (approximately 10%) (Applebaum, 1966). It is assumed that as the distance increases, the strength of these relationships diminishes (Löffler, 1998). This study investigates the following hypothesis:

H1: The retailer sales are positively related to the average distance of customers from the primary area of influence to the store.

Population Structure

The population structure gauges whether the purchasing behaviors of the residents align with the preferences of the company's target consumers (Turhan et al., 2013) (Gonzalez-Benito B and Gonzalez-Benito J, 2005). These variables encompass a range of factors including demographic attributes, shopping patterns, and commuting times (Turhan et al., 2013). A more behavior-centric approach includes models that consider variables like shopping frequency, the maximum distance consumers are willing to travel to the store, and their residential and workplace locations. Some studies unify geodemographic data, generally obtained using census databases, with people’s lifestyles databases (O Malley et al., 1995). This study gauges whether population structure can improve models’ predictive power to guide store location choices with the following hypotheses:

H2a: The retailer sales are related to the household number in the primary influence area.

H2b: The retailer sales are related to the household residents in the primary influence area.

H2c: The retailer sales are related to the average household residents per house across the primary influence area.

H2d: The retailer sales are related to the average household residents per house in the primary influence area census sectors with registered customers.

Economic Factors

Although the category of economic factors may intersect with demographic characteristics, it is essential to delve deeper to understand the financial resources available for retail expenditure (Turhan et al., 2013). When considering purchasing power, the location and store formats must differentiate between consumers in emerging markets and those with higher incomes (D'Andrea et al., 2006). This involves an analysis of various variables, including household income, sources and distribution of income, consumer buying potential, households' propensity to allocate their expenditure at the store, per capita sales figures, mobility conditions, and housing arrangements (rented or owned). This study investigates the following hypothesis:

H3a: The retailer sales are related to the average income of households across the primary catchment area.

H3b: The retailer sales are related to the average income of households in the primary catchment area census sectors with registered customers.

Competition

The competition among stores exhibits a paradoxical relationship. Proximity among competitors implies overlapping areas of influence, potentially leading to adverse effects on the performance of individual stores. However, competition can exert a more profound positive influence on the individual performance of each store within the cluster. This positive effect neutralizes or surpasses the negative impact of competition through the spillover effect generated by the presence of a retail cluster (Teller et al., 2016).

This relationship is intricate, and further situations need to be considered. When many competitors are in a specific area, they become less distinguishable by consumers and fail to attract more consumers to the cluster, leading to pressure for lower prices (Hoch et al., 1995) (González-Benito et al., 2005). Furthermore, isolated stores have lower price sensitivity by consumers than stores close to each other. Although the spatial rivalry is greater between stores of the same format than between different store formats, implying a hierarchical organization of consumers when choosing a store (González-Benito et al., 2005). When defining new store locations, retailers usually observe the spatial distance between retail stores, the size and number of competitors, the relative strength of competitors, competitors' sales volume, purchase alternatives, and the saturation level of the shopping area (Turhan et al., 2013) (Yanine et al., 2019). This understanding underscores the significance of considering competition within the location selection process. Accordingly, this study investigates the following hypotheses:

H4a: The retailer sales are related to the number of competitors within a radius close to the store (400 meters).

H4b: The retailer sales are related to the distance to the nearest competitor (in meters along the road network).

Footfall

Store sales can be generally predicted by observing the footfall in front of the store. This density of passers-by helps to define the economic viability of a given commercial location (Graham et al., 2019). Commemorative dates and days of the week have an essential impact on consumption patterns (Parsons, 2001), including daily fluctuations (Berry et al., 2016). This study also considered the possibility of a working day or weekend to moderate the footfall effect on daily ticket sales. When coupled with demographic data, pedestrian flow has the potential to elucidate the temporal fluctuations in demand over the course of the day (Waddington et al., 2019). Therefore, we analyze the following hypothesis:

H5: The retailer sales are positively related to the footfall in front of the store.

Access

The choice of store location must consider ease of access. Customers prefer more convenient stores and drive short distances to reach the store (Swoboda et al., 2013). Retailers must ensure that customers find the store easily and quickly (Thang and Tan, 2003). Additionally, retailers should account for more factors, such as the presence of parking lots, vehicle traffic, sidewalk, street width, and store opening hours (Turhan et al., 2013). Stores near transport hubs have superior economic results as they offer better pedestrian access (Kang, 2016), especially for brands focused on customers with low income (Parente et al., 2012). Thus, this study investigates whether:

H6a: The retailer sales are positively related to the proximity to a transport hub.

H6b: The retailer sales are positively related to the store location on the way to a transportation hub.

Store Characteristics

Regarding store and image attributes, the commonly used location selection criteria are the atmosphere, the area in square meters, the size of the storefront, and the number of checkouts (Turhan et al., 2013). Larger stores are usually more attractive because they have a more significant assortment of products (Reigadinha et al., 2017), and the visibility of the store, both for pedestrians and passengers in means of transport, can be objectively incorporated into quantitative models (Wood and Tasker, 2008). Thus, the following hypotheses are established:

H7a: The retailer sales are positively related to the total sales area.

H7b: The retailer sales are positively related to the storefront size.

Performance

The commercial point is chosen by comparing the best-expected performance among the available alternatives. Performance is an aggregate measure of the effect of consumer buying behavior (Kumar and Karande, 2000). The return potential of a retail location can be measured in terms of sales volume, profits, market share, number of customers, or price elasticity. Therefore, all criteria considered when evaluating a new store will be guided by the ability to increase the expected return (Turhan et al., 2013). The methods used to guide decisions about opening new stores are typically based on multivariate regression and spatial interaction (Turhan et al., 2013) (Kumar and Karande, 2000).

Based on the previous theoretical consideration, the hypothesized framework is presented on Figure *1*.

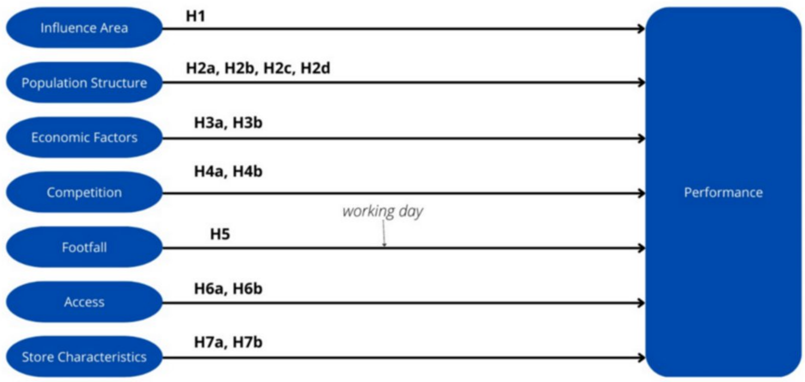

FIGURE 1: Framework for Store Location Effectiveness

Research Method

This research employed a quantitative, non-experimental case study design to analyze the effect of location selection criteria on sales forecasts. Data were collected from 22 stores of a Brazilian menswear retail chain located in commercial street centers. The study combined internal sales and customer data with external geospatial and demographic information. Using simple and multiple regression analyses, the research tested hypotheses and developed a predictive model to support location-based decision-making for new store openings.

The company is a men's clothing retailer in Brazil founded in 1987. The company has 38 small stores distributed in the state of São Paulo, located both in commercial street centers and shopping malls. The customer base is formed mainly by social classes B and C. The chain of stores fits into the specialized retailer category, offering products for a specific segment (men's fashion, mainly social), with a differentiation strategy (i.e., in terms of store layout and employee presentation) in typically small stores and a higher level of services, such as free tailoring and personal credit cards offered in partnership with financial companies.

The research was conducted in 22 stores of the chain located in commercial street centers. The reasons for this choice include the company's strategic expansion plan, preferably through new high street stores, and the essential differences between planned and unplanned retail centers, which required two specific sales forecast models.

To address the research problem, the specific approach was that of an observational, non-experimental study carried out via a single cross-sectional data collection, ensuring a numerical description of a population's trends, attitudes, or opinions when studying a sample (Johnson, 2001), (Creswell, 2017), and (Oehlert, 2010). Table *1* summarizes the variables considered in this study.

Criteria	Indicator	Hypothesis	Description	Reference
Performance	-	Independent variable	The number of daily and weekly tickets	(Graham et al., 2019)
Influence area: Extensions of the territory where the retailer or retail cluster customers are typically found	v01	H1: The retailer sales are positively related to the average distance of customers from the primary area of influence to the store.	The average distance of customers in the primary area of influence to the store	(Reigadinha et al., 2017), (Brown, 1993)
Population structure: Purchasing habits of the population residing in the area are congruent with the company's target consumer	v02	H2a: The retailer sales are related to the household number in the primary influence area.	Households throughout the primary area of influence	(Wang et al., 2015)
	v03	H2b: The retailer sales are related to the household residents in the primary influence area.	Residents throughout the primary influence area	(Shields and Kures, 2007), (Kumar and Karande, 2000)
	v04	H2c: The retailer sales are related to the average household residents per house across the primary influence area.	The average number of residents per household in the entire area of primary influence	(Shields and Kures, 2007)
	v05	H2d: The retailer sales are related to the average household residents per house in the primary influence area census sectors with registered customers.	The average number of residents per household in the census sectors of the primary area of influence with registered clients	(Shields and Kures, 2007)
Economic factors: Resources that can be spent on retail	v06	H3a: The retailer sales are related to the average income of households across the primary catchment area.	The average income of residents of the entire area of primary influence	(Shields and Kures, 2007)
	v07	H3b: The retailer sales are related to the average income of households in the primary catchment area census sectors with registered customers.	The average income of residents in the census sectors of the area of primary influence with registered clients	(Shields and Kures, 2007)
Competition: Competition between stores.	v08	H4a: The retailer sales are related to the number of competitors within a radius close to the store (400 meters).	Competitors within a radius of 400 meters	(Chen and Tsai, 2016), (Wang et al., 2015)
	v09	H4b: The retailer sales are related to the distance to the nearest competitor (in meters along the road network).	Distance from the nearest competitor (meters along the road network)	(Reigadinha et al., 2017)
Footfall: The density of passers-by in front of the store.	v10	H5: The retailer sales are positively related to the footfall in front of the store.	Daily/Weekly footfall	(Ho et al., 2013), (Kang, 2016); (Themido et al., 1998)
Access: Stores need to be ease to access. Customers prefer more convenient stores and	v11	H6a: The retailer sales are positively related to the proximity to a transport hub.	Distance to the nearest transport hub	(Kang, 2016)

drive short distances to reach the store.	v12	H6b: The retailer sales are positively related to the store location on the way to a transportation hub.	Way to transport hub (dummy)	(Chen and Tsai, 2016)
Store characteristics: The store and image attributes, such as the atmosphere, the area in square meters, the size of the storefront, and the number of checkouts	v13	H7a: The retailer sales are positively related to the total sales area.	Sales area (m^2)	(Chen and Tsai, 2016), (Li and Liu, 2012)
	v14	H7b: The retailer sales are positively related to the storefront size.	Storefront size (meters)	(Wang et al., 2015)

TABLE 1: Variables Description

Collecting data

The chosen performance metric for evaluating the sales potential of commercial outlets was the count of daily and weekly sales transactions, focusing on the stores' ability to convert passers-by into customers (Themido et al., 1998) (Graham et al., 2019). Data on the company's customer base were collected from September 16, 2020, to December 16, 2020, utilizing customer address records in transaction data from the company's enterprise resource planning system. The information accessed complied with the General Data Protection Law of 2018 and included customer addresses and their respective branch affiliations. On average, each store had 174 registered addresses.

In terms of the primary influence area, the study examines the average variable distance of customers concerning the store (v.01). The significance of the variable demonstrates the impact of geographic dispersion on store performance. Determining the influence area advanced the comprehension of the commercial hub's characteristics, such as its attractiveness and the impact on store performance. Influence areas were determined based on the proximity of customers to their residences: the primary influence area encompassed the region where 60% of the closest customers resided, followed by the secondary area comprising the next 25%, and the tertiary area including the remaining 15% of more distant customers. To determine the primary influence area of each store, two methods were employed: convex envelopes and isocots relative to the store's location. The dimensions and configurations (in the case of convex envelopes) of the primary influence area varied according to the geographic dispersion of 60% of each store's customer base (Applebaum, 1966). Geospatial data for all registered customers were processed using GIS software (QGis) to map the areas defined by these distances.

Regarding the population structure and economic factors, the household addresses of customers of the retailer stores were crossed with data from the 2010 Brazilian Institute of Geography and Statistics (IBGE, 2011), Demographic Census, grouped by census sector, to infer the characteristics of the customers. This crossing made it possible to ascertain the total number of households (v.02), the total population (v.03) of the area of primary influence, the average number of residents (v.04 and v.05), the monthly income of those responsible by households (v.06 and v.07), both in relation to all census sectors in the area of primary influence and the census sectors in the area of primary influence with registered clients.

The presence of competitors around the stores was inserted in two ways: the number of competitors within a radius of 400 meters (v.08) and the distance from the nearest competitor (v.09). The chain has more than one store within a 400-meter radius at two locations. They were considered in the model as competitors because of their potential for “cannibalisation” (Dibb and Simkin, 1994). The retailer is a comparative purchaser and benefits from the area’s attractiveness; therefore, it tends to be more affected by the micro-location and accessibility of the store than other stores (Li and Liu, 2012). Accordingly, the number of competitors in an area close to the store was examined, which was always within the limits of the size of the commercial centre itself (Chen and Tsai, 2016).

The footfall variable is employed to help predict performance, and as it is measured in individual terms, it was considered valid to adopt the same criterion for the dependent variable. This variable was obtained by collecting the number of pedestrians passing on the sidewalk in front of the store (v. 10) through manual counts made by the staff members of each store. Counting took place in all 22 stores at counting intervals of 10 minutes every hour. The period of collection was one week, from 10/26/2020 to 10/31/2020.

Modern alternatives to people counting, such as cell phone signal collection or people detection cameras (Graham et al., 2019), were considered in this study but were discarded due to the cost and retailer context, as stores often have street vendors stationed on the sidewalk in front of the store. The store’s sales team would usually move to the sidewalk to address customers. In the case of cell phones, the consulted companies stated that the range of coverage is wide (around 100 meters) and can result in different efficiencies depending on the location.

From the counts in 10-minute periods, the volume of passers-by was expanded to predict the number of pedestrians for periods of one hour. The count always took place in the middle of the hour that was intended to be predicted. This technique was performed following the instructions of previous studies (Davis et al., 1988), whereby a calculation formula was applied for pedestrian expansion that guaranteed an R^2 of 0.86 and a standard error of 0.18 in their study. To ensure that branch teams performed the counts similarly, a counting manual was developed in an appropriate language, informing about times and places for counting and the types of people to be counted. The stores received a table to fill in the counts and a manual flow counter.

The presence of transport hubs facilitates consumers' access to stores, especially if this access is the predominant type for the consumer profile (Parente et al., 2012) (Turhan et al., 2013) but can generate a volume of passers-by on the move with no intention to buy. These fundamentals were considered in this study in two ways: by analyzing the distance (through the road network) to the nearest transport hub (v.11) and determining whether the store is in a pedestrian crossing location (v.12), located on the same street as a transportation hub or perpendicular to the main avenue that contains a hub. This last variable was a dummy inserted into the model. Regarding the store's characteristics, the store's sales area (v.13) was considered as the total area minus the area for employees and stock storage. The storefront size (v.14) was measured in linear meters, which would help to demonstrate the store's visibility. Both data were obtained internally at the company.

Although other aspects unrelated to the location can also impact the performance of stores, this study chose to limit its scope to the location selection criteria. Salespeople accompany customers throughout the process. Thus, the salesperson's efficiency level may vary, as attitude, skills, personal factors, and motivation impact the sales team's performance (Churchill et al., 1985). Moreover, higher staff turnover differs between stores and is associated with a reduction in the store's performance, as measured by service level or profit margin (Ton and Huckman, 2008). Variables related to the refurbishment pattern of each store, the layout, and the investment in marketing were not considered in this study as all units follow a similar pattern and have a similar level of service. Aspects, such as the weather, can explain part of the variance of historical sales by retailers (Badorf and Hoberg, 2020), especially the incidence of rain reducing the effect on the traffic of people in commercial centers (Parsons, 2001).

In this study, the possibility that the flow of passers-by was influenced by the observation of whether the collection day was useful or on the weekend was considered, especially in poles where there is a large flow of pedestrians composed of workers moving during the week. Therefore, the business day dummy variable (v.15) was also collected.

Data analysis

The measured independent variables, correlating with the tickets sold at the store (dependent variable), served as the basis for creating an equation with predictive capacity for new stores. Previous studies used this methodology with practical success (Themido et al., 1998) in a network of gas stations in Portugal. Otherwise, the Simple Linear Regression method was preferred to validate the hypotheses between each explanatory and response variable.

Analysis concerning missing values, standardization of data, the treatment of outliers, the tests to verify the normality and homoscedasticity of the residues, the investigation of the coefficients of determination, the tests of individual and global significance, and tests to confirm possible multicollinearity were performed to guarantee the effectiveness of the regression models presented. The type of regression performed was stepwise, in the same way as implemented by previous studies (Simkin, 1989) (Themido et al., 1998). These studies employed a model with weekly tickets as the dependent variable, under the assumed moderation of the variable-business day.

The method starts the regression by the intercept, sequentially adding to the model the predictor variables that best improve the fit or by removing terms with low significance. In computational terms, it is agile because it performs more restricted searches, presenting lower variance but more bias (Hastie et al., 2009). This method is a good alternative for this study, which seeks to explore several variables that impact the location choice.

Results

The data collected in this study were standardized, and all variables were subtracted by the mean and divided by the standard deviation. Consequently, the predictors became comparable, and the interpretation of regression coefficients was facilitated (Schielzeth, 2010). In this study, observations with standard deviations greater than 3 or less than -3 were considered outliers (Osborne and Overbay, 2004). There was an observation above 3 for the Sales Area variable (v.13) of branch 47 (3.37) because this branch covered an extra 100 m² in comparison to the second largest store (the network's average sales area is 118 m²). In this case, the winsorization procedure was performed with the outliers, replacing the data of variables above the 95th percentile with the 95th percentile and below the 5th percentile with the 5th percentile (Ghosh and Vogt, 2012). Despite the adoption of the procedure, justifiable from the marketing perspective, the variable was not made a part of the predictive models, shown in (Table 2).

Variable	N	Constant	Coefficient	T-value	p-value	R²
Distance Average Customers (v.01)	22	0	-0.134	-0.607	0.5508	0.02
Households Convex Hull (v.02)	22	0	-0.209	-0.954	0.3515	0.04
Households Radius (v.02b)	22	0	-0.237	-1.092	0.2877	0.06
Residents Convex Hull (v.03)	22	0	-0.196	-0.894	0.3818	0.04
Residents Radius (v.03b)	22	0	-0.225	-1.033	0.3138	0.05
Average Residents Convex Hull (v.04)	22	0	0.452	2.266	0.0347*	0.2
Average Residents Radius (v.04b)	22	0	0.335	1.592	0.127	0.11
Average Residents Customers (v.05)	22	0	0.61	3.447	0.0026**	0.37
Income Convex Hull (v.06)	22	0	-0.443	-2.211	0.0388*	0.2
Income Radius (v.06b)	22	0	-0.313	-1.472	0.1567	0.1
Income Customers Primary (v.07)	22	0	-0.489	-2.51	0.0208*	0.24
Competitors Radius 400m (v.08)	22	0	-0.177	-0.805	0.4305	0.03
Distance from Competitor (v.09)	22	0	0.081	0.365	0.719	0.01
Footfall Weekly (v.10)	22	0	0.746	5.01	0.001**	0.56
Transport Hub Distance (v.11)	22	0	0.233	1.073	0.2959	0.05
Hub Passagem (v.12)	22	-0.033	0.122	0.248	0.8066	0
Sales Area (v.13)	22	0	0.292	1.399	0.1771	0.09
Storefront (v.14)	22	0	0.246	1.137	0.269	0.06

TABLE 2: Linear Regression Results

Note: * The correlation is significant at the 0.05 level (2 ends). ** The correlation is significant at the 0.01 level (2 ends).

To integrate the final equation, the stepwise regression process indicated the variables with p-values lower than 0.05 in each ticket prediction model sold. The variables related to the primary area of influence obtained through the isocots radius (v.02b, v.03b, v.04b, v.06b) were disregarded in the development of the models, as the simple regressions indicated that the variables formulated through the design area of influence in a convex wrapper (Convex Hull) showed a higher correlation with tickets sold. Thus, among the 14 independent variables initially considered in the weekly model, two of them made up the predictive model (Equation 1).

Weekly Tickets = 0.000 + 0.833 Footfall Weekly (v.10) + 0.419 Storefront (v.14) (1)

The residuals of the weekly predictive model were normal. Kolmogorov-Smirnov and Shapiro-Wilk normality tests were performed concerning the distribution of residuals. In both, the null hypothesis of population normality is rejected. In the Shapiro-Wilk test, a correlation coefficient is 0.968, with a p-value of 0.665, which also occurred in the Kolmogorov-Smirnov test, with a correlation coefficient of 0.160 and a p-value of 0.147.

The model's coefficient of determination (R^2) (weekly) reached convincing levels (69.6%), with a standard error of 0.551, indicating the satisfactory ability of the model to explain the variation in tickets sold (y). This coefficient of determination is in line with previous studies on the subject (Simkin, 1989) (Themido et al., 1998). The F test was performed and signified that the independent variables that made up the weekly model, $F(1.2) = 25.00$, $p < 0.001$, have a significant relationship with the response variable. T-tests were performed to verify the significance of each independent variable included in the model, and satisfactory results were obtained for both models, with p-values lower than 0.05. The occurrence of multicollinearity between the independent variables was also analyzed. From the measurement of Variance Inflation

Factors (VIFs), it was observed that the model's highest value found, 1.045 (Footfall Weekly and Storefront), signifies a low correlation between the independent variables, as shown in Table 3.

	Coefficient	Standard Error	Confidence Interval (95%)	T-value	P-value	VIF
Constant	0.000	0.118	(-0.246; 0.246)	0.000	1.000	
Footfall Weeks (v.10)	0.833	0.123	(0.575; 1.090)	6.769	0.000	1.045
Storefront (v.14)	0.419	0.123	(0.162; 0.677)	3.406	0.003	1.045

TABLE 3: Coefficient, Standard Error, Confidence Interval, p-value, and VIF

VIF, Variance Inflation Factor

The generated model had sufficient explanatory power to forecast new store sales, according to Table 4, with a Mean Absolute Percentage Error (MAPE) of 16.1% for this sample.

Store	Standardized Accomplished Tickets	Standardized Predicted Tickets	Residuals	Transacted Tickets	Expected Tickets	Residuals	% Error
Store 03	-1.98	-1.53	0.45	75	102.52	-27.52	37%
Store 04	-0.28	-0.61	-0.33	180	159.51	20.49	-11%
Store 05	0.37	-0.04	-0.41	220	194.74	25.26	-11%
Store 09	1.28	0.68	-0.60	276	238.77	37.23	-13%
Store 14	1.46	1.27	-0.19	287	275.35	11.65	-4%
Store 17	0.78	0.42	-0.36	245	222.86	22.14	-9%
Store 21	0.31	0.46	0.15	216	225.30	-9.30	4%
Store 26	1.10	0.81	-0.29	265	247.28	17.72	-7%
Store 27	1.04	0.51	-0.53	261	228.53	32.47	-12%
Store 28	0.32	-0.72	-1.04	217	152.74	64.26	-30%
Store 29	0.23	0.07	-0.15	211	201.56	9.44	-4%
Store 40	0.65	1.16	0.51	237	268.57	-31.57	13%
Store 41	-1.51	-0.71	0.80	104	153.43	-49.43	48%
Store 42	0.27	-0.22	-0.50	214	183.3	30.70	-14%
Store 44	0.40	0.60	0.20	222	234.32	-12.32	6%
Store 45	-0.78	-0.99	-0.21	149	136.23	12.77	-9%

Store 46	0.26	1.26	1.00	213	274.86	-61.86	29%
Store 47	0.44	0.23	-0.21	224	210.95	13.05	-6%
Store 51	-1.99	-1.64	0.36	74	96.11	-22.11	30%
Store 52	-1.26	-0.85	0.41	119	144.32	-25.32	21%
Store 53	-0.58	-0.49	0.1	161	167.10	-6.10	4%
Store 55	-0.52	0.32	0.84	165	216.65	-51.65	31%
						MAPE	16.1%

TABLE 4: Tickets Forecast, Tickets Accomplished, and Error Percentage of the Weekly Model by Store

MAPE, Mean Absolute Percentage Error

Based on ticket modeling, this precision is satisfactory for the managerial decision to select commercial locations, as already indicated by the low standard error of the model (0.551). The most distant prediction from reality (store 41) comes from a store located in a region of the city that in the past was home to a retail commercial hub, as well as many offices and company headquarters. Over the years, the number of stores has reduced, maybe because the motivation of passers-by going to work overlaps those motivated by shopping. As a result, the model overestimated the store's potential, projecting sales to be 48% higher.

Another discrepant forecast (store 51) points to an expectation of tickets sold 30% higher than that realized by the store. However, this branch is the only one in the entire study that has not yet reached maturity, having less than three months of activity at the time of the study. Hence, the equation can be an indication of the sales potential of this unit. Previous studies showed a similar level of error in predictions. Previous research (Simkin, 1989) claims that the models generated average errors of less than 10%, whereas the average errors were 15% and 17% in other model (Themido et al., 1998).

Discussion

The design of the influence area of the stores presented the demographic characteristics of the consumers and the distances travelled by the customers to reach the store. However, this distance was not validated as a determinant for the sales projection or the regression equation. Thus, hypothesis H1 did not show statistical significance. The same occurred for hypotheses H2a and H2b, which deal with the number of residents and dwellings in the influence area. Each commercial hub has different attractiveness characteristics. The commercial hub where store 53 is located has the characteristic of attracting customers from further distances, including from other states as the hub also has stores focused on wholesale trade. Consequently, the areas of influence are extensive. This phenomenon also occurs in stores 17, 29, and 52.

Hubs in cities in Greater São Paulo essentially attract the population closest to the store's surroundings, such as store 04. The positive sign of the coefficient denotes that the larger the store frontage, the higher the level of tickets sold when analyzed together with the weekly footfall variable (v.10). This finding can be attributed to two possible explanations: (1) store size is a measure of attractiveness. The larger it is, the wider the assortment of products offered and its gravitational attraction power (Reigadinha et al., 2017); (2) The size of the facade increases the visibility of the store and is correlated with expected sales (Wood and Tasker, 2008). Store 27, located in the interior of the state of São Paulo, is part of a single large popular commercial center in the city, with the capacity to attract consumers from nearby cities, resulting in a wide primary influence area (11.18 km). Even with all these characteristics differentiating each commercial hub (and, consequently, each commercial point), the variables related to the delimitation of the influence area were not significant for the predictive model.

Possible reasons can be listed for the non-significance of the variables related to the distance between the stores and the customers' homes (v.01) and the number of residents and households in the influence area (v.02 and v.03). For ticket prediction, the source of the flow is less important than the intensity and demographic characteristics of the flow. Competitive interactions are occurring between hubs that are not foreseen in this model, so the demand from around hubs with high population density will not necessarily be destined exclusively for the hub. An example is a hub located in downtown São Paulo, where the company has stores with low tickets sold. Poles have different sizes, varying their power of attraction according to this factor, as per the Theory of Spatial Interaction (Brown, 1993). The consumption dynamics for comparative shopping stores diverges from the dynamics of the supermarket and convenience retail, wherein distance or travel time to the store is more important (i.e., consumers are expected to go to the most logistically convenient location) (Dolega et al., 2016). For the retailer studied, the origin of customers has no direct relationship with their proximity to the store. In some examples, as in stores 40 and 44, the hubs attract demand from more remote locations wherein the lower income levels of the population are more in line with the popular profile of the existing stores in the hub.

The average of residents in the primary influence area (v.04) showed a relationship with the tickets sold. The more people live in the customers' homes, the greater the number of tickets sold. This variable possibly points to a type of family organization that is according to the brand object of this study (i.e., more people residing per household). The chain's stores located on the outskirts of São Paulo are visited by customers who live in larger families, thereby achieving good results in terms of tickets sold. While two of the company's stores, located in downtown São Paulo, whose homes comprised the fewest residents, are among the stores with the lowest sales performance. The number of residents signaled an ability to capture a consumer's behavioral profile through how families develop their homes, confirming H2c.

The average of residents only in the primary influence area census sectors, where the company has registered customers (v.05), also showed significance, validating hypothesis H2d. The coefficient of determination of this variable (37.3%) is higher than the variable of hypothesis H2c (20.4%), which reinforces the understanding that the number of residents is related to the forecast of tickets sold because when the analysis uses the census sector rather than the influence area, a stronger relationship is found. The result agrees with previous location studies, which observed the hypothesis that the number of residents per household in the influence area could positively affect sales if aligned with retailer segmentation (Shields and Kures, 2007). Income is another sociodemographic aspect of consumers, having the ability to explain expected sales. The income of the primary influence area revealed an R^2 of 19.6%. Conversely, the income of the census sectors with registered clients had an R^2 of 24%. Both statistical significances confirm hypotheses H3a and H3b.

The retailer maintains stores exclusively in popular commercial centers, so that there could be insufficient variability in the consuming public in the places studied, enabling the simple linear regression to indicate a relationship with sales. However, the findings were satisfactory and clarified both the importance of income and the type of income of the public at the hub that can allow the company to generate more tickets. The negative coefficients highlight that the lower the public's income, the greater the company's sales potential (i.e., -0.443 when we look at the entire area of primary influence, and -0.489 when we study only the census sectors with registered customers). It was not possible to validate the hypotheses related to competition. The distance from the nearest current (v.09) and variable competitors within a radius of 400 meters (v.08) presented high p-values, showing no evidence that hypotheses H4a and H4b are true.

The literature advocates that the proximity of competitors reduces consumers' price sensitivity (Hoch et al., 1995) and that the overlap of influence area and competition for the same market tends to reduce sales (Gonzalez-Benito et al., 2005). However, the positive effect of the agglomeration of stores can overcome the negative effect of competition (Teller et al., 2016). This study failed to justify any of these statements. A reason for the non-statistical significance of the variables can be determined by the impact that the size and attractiveness of commercial centers have on competition. In this sense, larger commercial centers can generate greater total demand, meaning that the presence of more competitors alone does not imply a reduced number of tickets for the retailer.

When analyzing a commercial location, visits are still necessary to learn about the nature of the location and its place in a broader competitive context (Wood and Tasker, 2008). Footfall (v.10) proved to be the location selection criterion with the highest coefficient both in the individual regression (0.557) and in the weekly predictive model (0.833). It was essential in determining the performance of a potential store, was consistent with the literature (Graham et al., 2019) (Wood and Browne, 2007) and helped validate H5. Additionally, the coefficient of determination of 55.7% in the simple linear regression reveals that this variable alone has the potential to explain a high portion of store sales. The chain stores with the highest sales invariably have an increased number of passers-by in front of the store. However, the literature recommends that flow is not enough. The other conditions must remain constant, such as store windows, assortment, environment, and level of service (Graham et al., 2019).

A reason for the preponderance of this variable is that smaller stores, as well as comparative shopping stores, do not generally exert a significant influence on determining the flow and power of attraction of people to the commercial hub, as they tend to take advantage of the existing flow (Wood and Tasker, 2008) (Graham et al., 2019). Furthermore, the decision to enter a clothing store is influenced by the communication and presentation of products that the windows provide (Sen et al., 2002). As the store impacts more people, they are more likely to make purchase decisions.

This research was also unable to infer that the store's proximity to a transport hub (v.12) is reflected in the tickets sold. Likewise, the distance from the store to the nearest transport terminal (v.11) presented an unsatisfactory p-value to characterize a relationship with the tickets sold. Hypotheses H6a and H6b were rejected. Street commercial centers develop in places where it is possible to take advantage of the high population density of urban centers. These places tend to have a more relevant presence of public transport options (Deliberali et al., 2019). When a store is located near a station or transportation terminal, it is expected to reap the benefits of the store's greater exposure and the greater volume of people impacted. However, this study indicates that for the company, there is no direct relationship between the presence of transport hubs and expected sales. Possibly, the flow of passers-by in front of the store is a better variable to infer the sales potential, even if part of this flow has been brought to the store due to the infrastructure of transport hubs, which act as facilitators of people's access to the retail hub. Even with these unpromising results, it is necessary to evaluate new commercial points, make a due observation of the means of transport available, determine if access for consumers will be facilitated and analyze the predominant type of displacement of passers-by, especially when a part is not motivated for shopping but focused on the simple displacement of working pedestrians through a shopping center.

Sales area (v.13) was not significant in the individual regression as it did not integrate the predictive model; hence, H7a was not validated. These results indicate that the available internal area does not impact the tickets sold due to the inconsistency in the formats of the properties, wherein some stores with more depth signify that a good part remains underused. In this way, it is more important to have a large storefront with an abundant internal area. The retailer's business model can have a defining role, as the market appeal is segmented, and the range of products offered is limited so that more sales area does not necessarily imply more assortment available to the consumer. Just as sales carried out by salespeople (and not self-service) can have an influence, space limitations can be overcome by the team.

Storefront (v.14) for not having demonstrated statistical significance in the simple linear regression (p-value = 0.2690) implied the rejection of H7b. However, this variable integrated the formation of the weekly predictive model, with a coefficient of 0.419. The positive sign of the coefficient indicates that the larger the storefront, the higher the level of tickets sold when analyzed together with the weekly footfall variable (v.10). This finding can be attributed to two possible explanations. First, store size is a measure of attractiveness. The larger it is, the wider the assortment of products offered and the greater its gravitational attraction power (Reigadinha et al., 2017). Second, the storefront increases store visibility and correlates with expected sales (Wood and Tasker, 2008).

Conclusions

This article evaluated the impacts of location selection criteria on the forecast of tickets sold in new stores. The generated equation presented satisfactory errors, and the linear regressions performed for each of the hypotheses could ascertain the impact of each location selection criteria on the tickets sold by the stores. The same approach can be replicated for other comparative purchase retailers by making eventual adjustments to the particularities of the businesses. The predictive model generated can be utilized in opportunities to open new stores, especially for small and medium businesses.

Based on the findings of this study, a company can carry out more informed analyses. First, there is the potential to use the model to predict outcomes based on the location selection criteria, widely discussed in the study, reducing the uncertainties associated with the opening of new stores. Similarly, it is possible to manipulate variables to carry out sensitivity analyses to understand the consequences of specific changes, such as the exchange of commercial points within the same hub. Additionally, it is possible to compare existing stores' expected and actual results to identify units performing above or below expectations and act when applicable. Lastly, this study allows the company to better prepare for the evaluation of the occupancy costs adequacy of current and target properties with the potential for ticket sales.

Regarding the location selection criteria, this paper allows a better understanding of each one in the sales forecast. As the flow of passers-by in front of the store is the most significant indicator of the sales potential, measurement and analysis become fundamental for the excellent selection of commercial points. Notably, it becomes less important to assess whether the hub's consumers come from long distances if there is sufficient volume to reach a satisfactory level of sales. When selecting a new commercial hub to open a store, it is essential to initially know the customers' origin, understand the income profile and the pattern of residents per household and whether these are aligned with the sociodemographic profile of the brand's customers. In the case of the company analyzed in this research, the ideal commercial centers are frequented by people of low income and with large family compositions. With regards to the store, new choices should favor larger storefronts, which guarantee more visibility. The store area should not be a deciding factor, if it meets the minimum requirements to ensure customers' convenience and be sufficient for the display and storage of products.

Despite having their hypotheses rejected in this study, some location selection criteria should be subject to managerial scrutiny before the opening of new stores. In this area, the intensity of the competitive dynamics of the commercial hub, even if it has not been validated in the linear regression, is crucial for the positioning of the brand and can result in performance below expectations if the model is applied in hubs with excess supply and overlap of influence areas. The tools for location assessment must be used in conjunction with the tacit knowledge of managers so that they can identify aspects through field visits that are difficult to be included in formal modeling, thereby increasing the effectiveness of choices.

This study did not analyze features such as product assortment, sales force efficiency, and the layout of a potential new store. Another limitation concerns the store's life cycle, as the sales predicted by the model quantify expectations for the new store's maturity, not for the initial months of operation. Additionally, the study collected the number of passers-by and tickets sold. However, it did not thoroughly investigate the ability of stores to convert passers-by into entrants and convert entrants into shoppers. However, as explained in the methodology section, the count of passers-by is susceptible to operational failures, even when care is taken to minimize errors. New studies can understand the impact of competition on the performance of a given shopkeeper in a commercial hub, differentiating the positive effects caused by the agglomeration of stores from the adverse effects of saturation. It is also suggested to study, from the outlook of a shopkeeper, the influence of the attractiveness of the commercial centers on the expected results, especially in contexts of overlaps between poles.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Carlos E. Lourenco, Adrian K. Cernev, Vanessa M. Santos, Leonardo F. Ferraz

Drafting of the manuscript: Carlos E. Lourenco, Adrian K. Cernev, Vanessa M. Santos, Leonardo F. Ferraz

Critical review of the manuscript for important intellectual content: Carlos E. Lourenco, Adrian K. Cernev, Vanessa M. Santos

Supervision: Carlos E. Lourenco

Acquisition, analysis, or interpretation of data: Leonardo F. Ferraz

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Acknowledgements

The authors would like to thank the Getulio Vargas Foundation for partially supporting this research. Leandro F. Ferraz is responsible for the data used in this research and is available upon request. The authors will keep the data for 5 years since data collection date. Leonardo Ferraz declares non-financial support from Dorinhos. Leonardo Ferraz is a board member of the retailer where the study was conducted. Carlos Lourenco was the advisor for his Master's thesis. The men's clothing retailer's advisory board consents to the use of the data.

References

1. Adeniyi O, Brown A, Whysall P: Retail location preferences: A comparative analysis. Journal of Retailing and Consumer Services. 2020, 55:102146. 10.1016/j.jretconser.2020.102146
2. Alaali A: Artificial intelligence integration and sustainable operations in physical retail stores: A comparative analysis of brands and SDG alignment. Business Development via AI and Digitalization. Hamdan A, Harraf A (ed): Springer, Cham; 2025. 538:71-87. 10.1007/978-3-031-62102-4_6
3. Alexander B, Varley R: Retail futures: Customer experience, phygital retailing, and the Experiential Retail Territories perspective. Journal of Retailing and Consumer Services. 2025, 82:104108. 10.1016/j.jretconser.2024.104108
4. Applebaum W: Methods for determining store trade areas, market penetration, and potential sales. Journal of Marketing Research. 1966, 3:127-41. 10.1177/002224376600300202
5. Badorf F, Hoberg K: The impact of daily weather on retail sales: An empirical study in brick-and-mortar stores. Journal of Retailing and Consumer Services. 2020, 52:101921. 10.1016/j.jretconser.2019.101921

6. Berry T, Newing A, Davies D, et al.: Using workplace population statistics to understand retail store performance. The International Review of Retail Distribution and Consumer Research. 2016, 26:375-95. 10.1080/09593969.2016.1170066
7. Brown S: Retail location theory: evolution and evaluation. The International Review of Retail Distribution and Consumer Research. 1993, 3:185-229. 10.1080/09593969300000014
8. Chen LF, Tsai CT: Data mining framework based on rough set theory to improve location selection decisions: A case study of a restaurant chain. Tourism Management. 2016, 53:197-206. 10.1016/j.tourman.2015.10.001
9. Creswell JW, Creswell JD: Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Sage Publications, New York; 2017.
10. Churchill GA, Ford NM, Hartley SW, et al.: The determinants of salesperson performance: A meta-analysis. Journal of Marketing Research. 1985, 22:103-18. 10.1177/002224378502200201
11. Davis SE, King LE, Robertson HD: Predicting pedestrian crosswalk volumes. Transportation Research Record. 1988, 1168:25-30.
12. Deliberali E, Brandão M, Bizarrias F: Repurchase intention conditioned to economic confidence in main street retail districts. Brazilian Business Review. 2019, 16:589-606. 10.15728/bbr.2019.16.6.4
13. Dolega L, Pavlis M, Singleton A: Estimating attractiveness, hierarchy and catchment area extents for a national set of retail centre agglomerations. Journal of Retailing and Consumer Services. 2016, 28:78-90. 10.1016/j.jretconser.2015.08.013
14. D'Andrea G, Lopez-Aleman B, Stengel A: Why small retailers endure in Latin America. International Journal of Retail & Distribution Management. 2006, 34:661-73. 10.1108/09590550610683184
15. Dibb S, Simkin L: Retail store performance model the impact of own-branch ‘friendly fire’. International Transactions in Operational Research. 1994, 1:479-87. 10.1111/j.1475-3995.1994.00324.x
16. Ghosh D, Vogt A: Outliers: an evaluation of methodologies. Joint Statistical Meetings. 2012, 3455-460.
17. Gonzalez-Benito B, Gonzlez-Benito J: The role of geodemographic segmentation in retail location strategy. International Journal of Market Research. 2005, 47:295-316. 10.1177/147078530504700305
18. González-Benito O, Muñoz-Gallego PA, Kopalle PK: Asymmetric competition in retail store formats: Evaluating inter- and intra-format spatial effects. Journal of Retailing. 2005, 81:59-73. 10.1016/j.jretai.2005.01.004
19. Graham C, Khan K, Ilyas M: Estimating the value of passing trade from pedestrian density. Journal of Retailing and Consumer Services. 2019, 46:103-11. 10.1016/j.jretconser.2017.10.005
20. Hastie T, Tibshirani R, Friedman J: The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, New York; 2009.
21. Hensel R, Visser R, Overdiek A, et al.: A small independent retailer's performance: Influenced by innovative strategic decision-making skills?. Journal of Innovation & Knowledge. 2021, 6:280-89. 10.1016/j.jik.2021.10.002
22. Ho HP, Chang CT, Ku CY: On the location selection problem using analytic hierarchy process and multi-choice goal programming. International Journal of Systems Science. 2013, 44:94-108. 10.1080/00207721.2011.581397
23. Hoch SJ, Kim BD, Montgomery AL, et al.: Determinants of store-level price elasticity. Journal of Marketing Research. 1995, 32:17-29. 10.1177/002224379503200104
24. IBGE. (2011). Accessed: February 4, 2021: http://ftp.ibge.gov.br/Censos/Censo_Demografico_2010/Resultados_do_Universo/Agregados_por_Setores_Censitarios/.
25. Johnson B: Toward a new classification of nonexperimental quantitative research. Educational Researcher. 2001, 30:3-13. 10.3102/0013189x030002003
26. Kang CD: Spatial access to pedestrians and retail sales in Seoul, Korea. Habitat International. 2016, 57:110-20. 10.1016/j.habitatint.2016.07.006
27. Kumar V, Karande K: The effect of retail store environment on retailer performance. Journal of Business Research. 2000, 49:167-81. 10.1016/s0148-2963(99)00005-3
28. Li Y, Liu L: Assessing the impact of retail location on store performance: A comparison of Wal-Mart and Kmart stores in Cincinnati. Applied Geography. 2012, 32:591-600. 10.1016/j.apgeog.2011.07.006
29. Litz RA, Rajaguru G: Does small store location matter? A test of three classic theories of retail location. Journal of Small Business & Entrepreneurship. 2008, 21:477-92. 10.1080/08276331.2008.10593436
30. Löffler G: Market areas - a methodological reflection on their boundaries. GeoJournal. 1998, 45:265-72. 10.1023/a:1006980420829
31. Mendes AB, Themido IH: Multi-outlet retail site location assessment. International Transactions in Operational Research. 2003, 11:1-18. 10.1111/j.1475-3995.2004.00436.x
32. NIELSENIQ. (2022). Accessed: December 3, 2023: https://nielseniq.com/global/pt/landing-page/ebit/nielseniq-ebit-brasil/webshoppers/.
33. Oehlert GWA: First Course in Design and Analysis of Experiments. University of Minnesota, 2010.
34. Osborne JW, Overbay A: The power of outliers (and why researchers should always check for them. Practical Assessment, Research, and Evaluation. 2004, 9:6. 10.7275/qf69-7k43
35. O Malley L, Patterson M, Evans M: Retailing applications of geodemographics. Marketing Intelligence & Planning. 1995, 13:29-35. 10.1108/02634509510083482
36. Parente J, Brandão M, Miotto A, Plutarco F: Main street retail districts or shopping centers? Comparing the preferences of low-income consumers. Brazilian Business Review. 2012, 154-79. 10.15728/bbrconf.2012.7
37. Parsons AG: The association between daily weather and daily shopping patterns. Australasian Marketing Journal. 2001, 9:78-84. 10.1016/s1441-3582(01)70177-2
38. Reigadinha T, Godinho P, Dias J: Portuguese food retailers - Exploring three classic theories of retail location. Journal of Retailing and Consumer Services. 2017, 34:102-16. 10.1016/j.jretconser.2016.09.015
39. Schielzeth H: Simple means to improve the interpretability of regression coefficients. Methods in Ecology and Evolution. 2010, 1:103-13. 10.1111/j.2041-210x.2010.00012.x
40. Sen S, Block LG, Chandran S: Window displays and consumer shopping decisions. Journal of Retailing and Consumer Services. 2002, 9:277-90. 10.1016/s0969-6989(01)00037-6
41. Seong EY, Lim Y, Choi CG: Why are convenience stores clustered? The reasons behind the clustering of similar shops and the effect of increased competition. Environment and Planning B Urban Analytics and City Science. 2021, 49:834-46. 10.1177/23998083211021870

42. Shaikh SA, Memon M, Kim KS: A multi-criteria decision-making approach for ideal business location identification. Applied Sciences. 2021, 11:4983. 10.3390/app11114983
43. Shields M, Kures M: Black out of the blue light: An analysis of Kmart store closing decisions. Journal of Retailing and Consumer Services. 2007, 14:259-268. 10.1016/j.jretconser.2006.07.007
44. Simkin L: SLAM: Store location assessment model—Theory and practice. Omega. 1989, 17:53-58. 10.1016/0305-0483(89)90020-0
45. Swoboda B, Berg B, Schramm-Klein H, Foscht T: The importance of retail brand equity and store accessibility for store loyalty in local competition. Journal of Retailing and Consumer Services. 2013, 20:251-262. 10.1016/j.jretconser.2013.01.011
46. Teller C, Wood S, Floh A: Adaptive resilience and the competition between retail and service agglomeration formats: an international perspective. Journal of Marketing Management. 2016, 32:1537-1561. 10.1080/0267257x.2016.1240705
47. Economist. (2021). Accessed: December 3, 2023: https://www.economist.com/business/2021/05/22/after-the-pandemic-boom-online-retail-sales-are-slowing.
48. Thang DCL, Tan BLB: Linking consumer perception to preference of retail stores: an empirical assessment of the multi-attributes of store image. Journal of Retailing and Consumer Services. 2003, 10:193-200. 10.1016/s0969-6989(02)00006-1
49. Themido IH, Quintino A, Leitão J: Modelling the retail sales of gasoline in a Portuguese metropolitan area. International Transactions in Operational Research. 1998, 5:89-102. 10.1111/j.1475-3995.1998.tb00106.x
50. Turhan G, Akalın M, Zehir C: Literature review on selection criteria of store location based on performance measures. Procedia - Social and Behavioral Sciences. 2013, 99:391-402. 10.1016/j.sbspro.2013.10.507
51. Ton Z, Huckman RS: Managing the impact of employee turnover on performance: the role of process conformance. Organization Science. 2008, 19:56-68. 10.1287/orsc.1070.0294
52. Waddington T, Clarke G, Clarke MC, Hood N, Newing A: Accounting for temporal demand variations in retail location models. Geographical Analysis. 2019, 51:426-47. 10.1111/gean.12179
53. Wood S, Tasker A: The importance of context in store forecasting: The site visit in retail location decision-making. Journal of Targeting Measurement and Analysis for Marketing. 2008, 16:139-55. 10.1057/jt.2008.3
54. Wood S, Browne S: Convenience store location planning and forecasting - a practical research agenda. International Journal of Retail & Distribution Management. 2007, 35:233-55. 10.1108/09590550710736184
55. Yanine F, Cordova FM, Valenzuela L, et al.: A fresh look to an old problem: saturation in the retail market, and how it affects both retailers and consumers. Indian Journal of Science and Technology. 2019, 12:1-10. 10.17485/ijst/2019/v12i18/144604
56. Zhang DJ, Dai H, Dong L, Wu Q, Guo L, Liu X: The value of pop-up stores on retailing platforms: evidence from a field experiment with Alibaba. Management Science. 2019, 65:5142-151. 10.1287/mnsc.2019.3410
57. Wang FF, Chen LF, Su CT: Location selection using fuzzy-connective-based aggregation networks: a case study of the food and beverage chain industry in Taiwan. Neural Computing and Applications. 2015, 26:161-70.