

Portuguese Food Retailers - Exploring three classic theories of retail location

Abstract

This article sets forth results from an exploratory empirical study that aimed to test the predictive validity of three classic location theories: central place theory, spatial interaction theory and the principle of minimum differentiation. Correlation, linear regression and analytical procedures in a Geographic Information System were used to reveal relationships between variables. The results show that all theories find significant support. It was possible to relate store location with distance to the centre, population density and competitor's location. It was also found that store location is related to consumer's proximity to the store (when stores are smaller) and the store's attractiveness to the consumer (when stores are larger).

Keywords: Retail Location; Food Retail; Portugal; Correlation Analysis; Classical Location Theories; Geographic Information System

1. Introduction

In Europe countries are still facing economic challenges with direct impact on consumer's confidence levels and family's spending. These economic challenges, combined with increasing consumer mobility, increasing electronic commerce, changing lifestyles along with a consumer much more diligent with his time and money, are forcing thousands of small stores, retail chains and shopping centres to rethink their strategy. As a fundamental component of national economy and an important element in organizing Portuguese sociocultural practices, retail, Portugal's GDP single most

contributor (INE 2013b), is becoming a priority to local and central administration. The retail sector has, since the 1990s, been the main player in the property market in Portugal. In fact, the retail gross leasable area (GLA) per 1.000 inhabitants in Portugal is higher than EU27 average and higher than UK or France retail GLA (Cushman&Wakefield 2012). The increase in the supply of retail space in Portugal has kept pace with this industry's success story, having recorded high growth rates up to 2009, when the market started to reveal a slower supply growth rate. The Portuguese market has been clearly dominated by institutional investors with international funds focusing on retail sectors and prime office space. The retail sector alone is responsible for 44% of the total property investment market between 1990 and 2013 (Cushman&Wakefield 2012).

Of all the elements of the retail mix, location is often referred to as the most important one (Brown 1994; Cox and Brittain 2004; Davies and Harris 1990; Ghosh 1994; Levy and Weitz 2009). Not only is the store's location generally equivalent to convenience and to the minimization of consumers' shopping efforts (Reimers and Clulow 2004), but it also involves the retailer in a long term commitment with the chosen location, given the amount of investment involved. Recently, a changing retail environment is augmenting the location importance as retail economic groups develop multi-outlets chains of small or medium stores. For the past decades the range of choices available to a retailer has continued to expand. Modern retail space (shopping centres, retail parks and factory outlet centres) in Portugal as at the end of 2013, amounted to more than 3.65 million square meters (Cushman&Wakefield 2012).

From the retail spatial patterns line of investigation, the four most important theories of retail location emerge: Central Place Theory (Christaller 1933), Spatial Interaction

Theory (Reilly 1929, 1931), Bid Rent Theory (Haig 1927) and the Principle of Minimum Differentiation (Hotelling 1929). Despite the theories' positivist and simplified premises, results of empirical investigations in this field show that the spatial patterns observed in reality are broadly supportive of the hypothesised patterns (Brown 1993, 1994; Litz and Rajaguru 2008). Since the bulk of empirical investigation is more than 30 years old [the case of the classic studies of Berry (1963) and Scott (1959)] and since retailing spatial structures has changed dramatically since then, the question is: can we still find support for the hypothesised patterns in an empirical investigation today? It is the main purpose of this investigation to conduct an exploratory research to test the predictive validity of three classic retail location theories. Due to the lack of reliable and comparable data on land values, Bid Rent Theory will not be tested.

Is the priority still on locating stores near the centre (Central Place Theory)? Is the store location power of attraction dependent on the distance to the consumers or the assortment carried by the store (Spatial Interaction Theory)? Do retailers still tend to cluster tightly together (Principle of Minimum Differentiation)? This study seeks to advance understanding concerning the relevance of classic theories of retail location for Portuguese food retailers. All main theories were formulated more than 30 years ago, based on a reality very different from the current Portuguese reality. Given the specific characteristics of this small country and the recent changes in the Portuguese retail environment it is crucial to understand if the classic theories of retail location are still relevant to the Portuguese retail players. Although international literature and empirical investigations about retail location are extensive and diversified, there's a deficit of academic work on the subject in Portugal. This investigation looks to fill, partially, these apparent research need. The findings of this research investigation will be important and especially relevant for large retailers, current and future, which manage

multiple stores and each year find themselves involved in store opening, closing and restructuring decisions. Also small retailers, without the power of attraction of large retailers, can benefit from these findings since their success relies heavily on the quality of their location. It can also be important for future professionals and entrepreneurs in the retail sector, in particular regarding their training. In another perspective it can also be relevant to the local and central government concerning spatial planning and land management.

After this brief introduction, in the next section we present a review of relevant literature and methodological approaches to the application of location theories. Next, we present the study area and data, followed by the methodology where we advance a set of hypotheses linking specific locational variables to the test of the predictive validity of three classical location theories. Finally we present, and discuss, the results and the main conclusions.

2. Theoretical and Methodological Framework

It was in the brief six year period between 1927 and 1933 that the four vital theories of our understanding of retailer's location were conceived. More than eighty years afterwards, Central Place Theory (Christaller 1933), Spatial Interaction Theory (Reilly 1929, 1931), Bid Rent Theory (Haig 1927) and the Principle of Minimum Differentiation (Hotelling 1929), still attract considerable attention from the academic world, but also some controversy (Brown 1993). As formulated by Christaller (1933), **Central Place Theory** seeks to explain the size, distribution and number of urban centres, and the hierarchical relationship among them, in an economy where consumers are uniformly distributed, with identical buying power and fully informed, and where sellers behave in rational, perfectly competitive and profit maximizing manner in a free

entry and equivalent costs market. Transport costs vary linearly with distance and the consumer, in order to minimize costs, acts rationally visiting the nearest centre offering the desired good or service (Brown 1993; Rodrigues, Vala and Monteiro 2002). Central Place Theory predicts that, given the increase in transport cost with increasing distance from the source of supply, i.e. from the centre, the demand for a particular good or service decreases. Beyond certain point, the good or service demand is zero and the distance consumers are willing to travel for the purchase of a specific good or service is called market area or range of a good (Brown 1993; Scott 1970). The larger the centre, the larger the market area and more establishments and types of businesses there are, providing a greater offer of goods and services than a small centre (Scott 1970). Using this approach Christaller demonstrates how the hierarchy of urban centres is formed. The centres at the top of the hierarchy have larger dimensions, namely in terms of population. Therefore, flows between centres occur only in vertical ascending direction, i.e., an urban centre of order n uses a higher order centre to search for a function it does not possess, belonging to the higher order centre trade area (Brown 1993; Mafra and Silva 2004; Rodrigues et al. 2002; Scott 1970). Unlike Central Place Theory that assumes that consumers patronize the nearest centre that offers the required good or service, **Spatial Interaction Theory** is based on the assumption that consumers trade off the attractiveness of alternative shopping areas against the deterrent effect of distance (Brown 1993). In other words, consumers may choose a less attractive shopping area closer to them or a distant centre with a larger offer of goods and services. Reilly's (1931) retail gravitation law, which states that in normal conditions two cities draw trade from an intermediate city approximately in direct proportion to their populations and in inverse proportion to the square of the distances of these two cities to the intermediate city, pioneered the gravitational models research (Brown 1993;

Lee and Pace 2005; McGoldrick 2002; Mendes 2005). These models, inspired by Newton's gravitational law, try to explain human behaviour related to retail activities through the gravitational phenomenon around big sized or nearest stores/cities. Several empirical tests were made to the gravitational postulates with mixed results (Brown 1993; Lee and Pace 2005). Although less well known than Central Place Theory or Spatial Interaction Theory, Hotelling's (1929) **Principle of Minimum Differentiation** laid the foundations for the study of micro-scale retail location (Brown 1994). Hotelling's model considers a linear market (e.g. street, road) in which demand is inelastic and identical, transport costs are constant, and consumers are evenly distributed and patronize stores solely on the basis of delivered prices (the price of the good plus transportation cost). If two retailers, selling the same type of product, decide to enter this market both will locate their store in the middle of the market and sell at the same price (Brown 1989; Hotelling 1929). If a third retailer enters this market he will locate near the other two retailers, but not between them, since he desires to attract as large a market as possible. If more and more retailers want to enter the market there will be a tendency for them to cluster (Hotelling 1929; Huang and Levinson 2008). Hotelling (1929) also maintained that the clustering phenomenon is visible on a plane as well as on a linear market when consumers are unevenly distributed. The introduction of uncertainty and risk reduction behaviour (i.e. the risk of consumers not finding what they are looking for and the risk of retailers being overtaken by competitors) into Hotelling's model by Webber, shows that agglomeration is the inevitable outcome (Brown 1989). Furthermore, the higher the variability of consumer demand, the higher the propensity to agglomerate. The introduction of these new variables also helped to formalize Nelson's (1958) Theory of Cumulative Attraction. From his extensive empirical surveys of consumer behaviour (Brown 1993), Nelson

proposes that a given number of stores dealing in the same merchandise will do more business if they are located adjacent or in proximity to each other than if they are widely scattered (Nelson 1958). Consumer's desire to compare the offerings of several stores before purchase, especially for items where price quality and fashion are important considerations (such as furniture, clothing and motor cars) led Nelson to his location law.

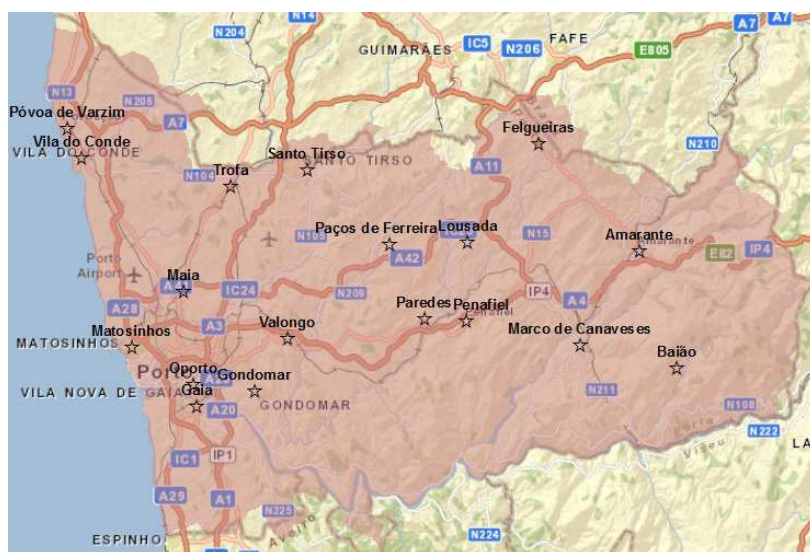
Since the 1920s, there has been a growing interest in the application of location theories to solve location decision problems, namely the decision to locate a new store. Spatial Interaction Theory has underpinned significant advances in this field. Various models for determining optimal store location, using several parameters like store performance, competition or economic factors, have been proposed [see Turhan, Akalin and Zehir (2013) for a review]. Most of these models use gravity modelling techniques (Benoit and Clarke 1997; Beule, Poel and Weghe 2014; González-Benito 2005; Lee and Pace 2005; Li and Liu 2012) or econometric techniques (Hymel 2014; Litz and Rajaguru 2008; Pennerstorfer and Weiss 2013; Themido, Quintino and Leitão 1998). Although the Central Place Theory and the Principle of Minimum Differentiation have been the focus of significant research activity, an even greater interest has been raised by Spatial Interaction Theory. Applications of Central Place Theory include, among others, shopping centres catchment areas and their position in retail hierarchy (Dennis, Marsland and Cockett 2002), urban and regional growth (Davies, Holz and Robertus 1999; Hsu 2012; Nakamura 2014) and relationship between centrality and economic development (Daniels 2007; Nogueira, Crocco, Figueiredo and Diniz 2014). The methods generally adopted were questionnaire surveys, econometric and mathematical modelling. Applications of Hotelling's Principle of Minimum Differentiation focus mainly on the clustering hypothesis (Borrell and Fernández-Villadangos 2010; Irmen

and Thisse 1998; Netz and Taylor 2002) and price competition (Barreda-Tarrazona, García-Gallego, Georgantzís, Andaluz-Funcia and Gil-Sanz 2011; Hausman and Leibtag 2007; Zhang 1995) using a similar methodology as other works related to Central Place Theory research. In recent years, the use of Geographic Information Systems (GIS) tools as a methodological option is becoming very popular among researchers (Cheng, Li and Yu 2007; Murad 2011; Önden, Sen and Sen 2012; Picone, Ridley and Zandbergen 2009; Porta et al. 2009; Roig-Tierno, Baviera-Puig, Buitrago-Vera and Mas-Verdu 2013; Shields and Kures 2007; Suárez-Vega, Santos-Peñate and Dorta-González 2012; Wang, Chen, Xiu and Zhang 2014). GIS tools have been used to support a wide range of research techniques (e.g. gravity modelling, statistical analysis, AHP) or as an independent research tool. Methodologies used for retail outlet location research have become more sophisticated as a result of modelling procedures brought about by GIS (Birkin, Clarke and Clarke 2002; Duggal 2007).

3. Study area and data

Oporto district is the second biggest district in population and one of the dominant districts in the Portuguese Retail scene. As shown in Figure 1, Oporto district has 18 municipalities. This region was selected because of the primary interest in examining retail stores in the urban area.

Figure 1 - Oporto District and Municipalities



The sample includes 273 stores, property of the eight main food retailers in Portugal (Table 1).

Table 1 - Stores profile

Number of Stores	Retail Chain	Operational format
48	Sonae SGPS, S.A.	Hypermarket, supermarket
69	Jerónimo Martins, SGPS, S.A.	Supermarket
94	Grupo Dia - Distribuidora Internacional de Alimentación, S.A. (Dia, S.A.)	Discount Store
6	Grupo Auchan Portugal, S.A.	Hypermarket, supermarket
35	Lidl & Cia	Discount Store
17	ITM Développement International (Intemarché)	Supermarket
2	Centres Distributeurs E. Leclerc	Supermarket
2	El Corte Inglés, S.A.	Department Store

Data acquired for this study include store and city hall information and data layers of Oporto district. The district demographic data were downloaded from the Portuguese Census Bureau website (INE 2013a). The store and city hall information, including addresses and store size, were collected from the retailer's and city hall's websites.

To test the hypotheses that will be formulated, the relationship between the following variables (Table 2) will be explored.

Table 2 - Variables Description

Dependent Variables	Description
Store size	Store's total size in square meters
Stores Density	Number of stores per parish area in square kilometre
Store's distance to the nearest competitor	Distance, in meters, to the nearest store from a different chain
Independent Variables	Description
Store's distance to the centre	Distance, in meters, between the store and city hall
Population density	Number of inhabitants per square kilometre

Store size is used here as a measure of attractiveness (larger stores and shopping centres are more attractive since they offer a larger assortment of goods). Stores density will be used as a measure for competitor density. The nearest competitor is considered to be the nearest store of any food retailer from a different chain. The centre is defined as the address of the city hall, from each municipality of Oporto district, since this building is usually inside or adjacent to the Central Business District (Cutsinger, Galster, Wolman, Hanson and Towns 2005; Galster et al. 2001; Pence 2011; Wolman et al. 2005). Population density is used as a measure of customer density.

4. Hypotheses and Methodology

This exploratory research examines the location pattern of food retailers in Oporto district, Portugal, to test the predictive validity of three classic location theories: central place theory, spatial interaction theory and the principle of minimum differentiation.

To explore the predictive validity of the classic theories, five hypotheses will be tested:

H1a. Store location is related to the distance to the centre (Central Place Theory)

H1b. Store location is related to the size of the centre (Central Place Theory)

In order to test the first two hypotheses, the relationship between store's distance from the centre and population density will be explored. According to Central Place Theory consumers act rationally by visiting the nearest centre that supply the merchandise, or service, required. The theory also predicts that demand for a good, or service, declines with distance from the source of supply, i.e. the centre. It is thus expected that the distance to the centre will be smaller in larger centres. The relationship between population density and stores density will also be explored since it is predicted that the larger the centre, namely in population, the larger the market area and the number of establishments. A positive relationship is expected.

H2a. Store location is related to the surrounding store density (Principle of Minimum Differentiation)

H2b. Store location and competitor's store location are related to the distance to the centre (Principle of Minimum Differentiation)

According to Hotelling's theory, as more and more sellers of the same commodity arise, the tendency is to cluster since they desire as large a market as possible, therefore the relationship between the distance to the nearest competitor and stores density will be explored. A negative relationship is expected, i.e. the larger the stores density the smaller the distance between competitors. The theory also predicts that firms, when clustered, tend to locate near the centre of the market area. To explore the validity of this premise the relationship between store's distance from the centre and the distance to the nearest competitor will be tested. A positive relationship is expected.

H3. Store location is related to consumer's proximity to the store or the store's attractiveness to the consumer (Spatial Interaction Theory)

The final hypothesis, concerning Spatial Interaction Theory, will test if the store's location is dependent on consumer's proximity to the store or the store's attractiveness to the consumer. As predicted by the theory, consumers trade off the attractiveness of alternative shopping areas against the deterrent effect of distance. In other words, consumers may bypass the nearest centre for a more distant but better equipped destination. In order to test this assumption, the relationship between store's size and population density will be explored with an expected negative relationship.

The emphasis is on understanding whether the Portuguese food retailers display different location preference from the perspective of centrality, agglomeration and equilibrium between store size and distance from the consumer.

Analytical procedures in GIS and statistical techniques, using IBM SPSS Statistics 20, have been applied to carry out the analysis in this study. All stores addresses were geocoded into ArcGIS and were submitted to distance calculations procedures, such as Near (Euclidean distance) and Average Nearest Neighbor, store density estimation, store pattern distribution and cluster analysis. Statistical techniques such as bivariate correlation and linear regression analysis were also applied.

5. Results and Discussion

The bivariate correlation results are presented in Table 3 and the linear regression results in Table 4.

Table 3 – Descriptive Statistics and Correlation Matrix

	Mean s.d.	1	2	3	4	5
1. Store size	1230.56 1832.17	1.0	-0.010	0.039	0.016	-0.166**
2. Stores Density	0.61 0.82		1.0	-0.638**	-0.484**	0.752**
3. Store's distance to the nearest competitor	1113.39 1498.43			1.0	0.421**	-0.344**
4. Store's distance to the centre	3849.1 3499.48				1.0	-0.202**
5. Population density	3,509.08 2,510.94					1.0

** . Correlation is significant at the 0.01 level (2-tailed).

Table 4 - Linear Regression Results

	1	2	3	4	5
Dependent Variable	Store's distance to the centre	Stores Density	Store's distance to the nearest competitor	Store's distance to the nearest competitor	Store size
Constant	5,181.2** (351.275)	-0.0206 (0.0885)	2,384.7** (265.948)	180.675 (111.465)	1,443.1** (190.479)
Independent Variable					
Population Density	-0.380** (0.081)	0.00027** (0.00003)			-0.061 (0.044)
Store's distance to the centre				0.242** (0.021)	
Stores density			-856.935** (261.232)		
Adjusted R²	0.071	0.484	0.091	0.318	0.003
F-Statistics	21.715**	92.137**	10.761**	127.685**	1.880

** . Rejection of null hypothesis at 0.01 level
Value in parentheses (.) are standard errors

Both correlation and linear regression results support Hypotheses 1a and 1b (Central Place Theory) as the expected negative relationship between store's distance to the centre and population density and the expected positive relationship between store density and population density are verified. Both statistical techniques also support Hypotheses 2a and 2b (Principle of Minimum Differentiation) as it was possible to establish a negative linear relationship between store's distance to the nearest competitor and stores density and a positive linear relationship between store's distance

to the nearest competitor and store's distance to the centre. The results only partially support Hypothesis 3 (Spatial Interaction Theory). Although the expected negative relationship between store size and population density is verified, the obtained coefficient is not significantly different from zero. However when extreme values are removed (12 stores with more than 3,087m²) a statistically significant relation between store size and population density is verified. The size of these 12 stores is very different from the majority of the sample because they correspond to the few hypermarkets of Oporto district.

To further test centrality and agglomeration of food retailers in Oporto district, some of ArcGIS spatial analysis and statistics tools will be used. The nearest neighbor index measures the distance between each store and its nearest store, and then calculates the average (Ebdon 1986). A smaller average distance (i.e., "observed distance") in comparison with a random distribution of the same number of stores (i.e., "expected distance") indicates a clustered pattern, and a larger average distance represents a dispersed pattern. Moreover, the nearest neighbor method reports a z-score to indicate whether the identified pattern is different from a random distribution in a statistically significant manner. The nearest neighbour index was calculated both for the whole district and also for each of the existing municipalities. It can be seen that the stores are located in a clustered way, much different from the random pattern. The observed distance is smaller than the expected distance and the ratio is less than 1. Given the z-score of -14.97 there is less than 0.1% likelihood that this clustered pattern could be the result of random chance. When the analysis is applied considering each municipality, all the results that are statistically significant indicate also a highly clustered pattern. These results lend strong support for the Principle of Minimum Differentiation. Table 5

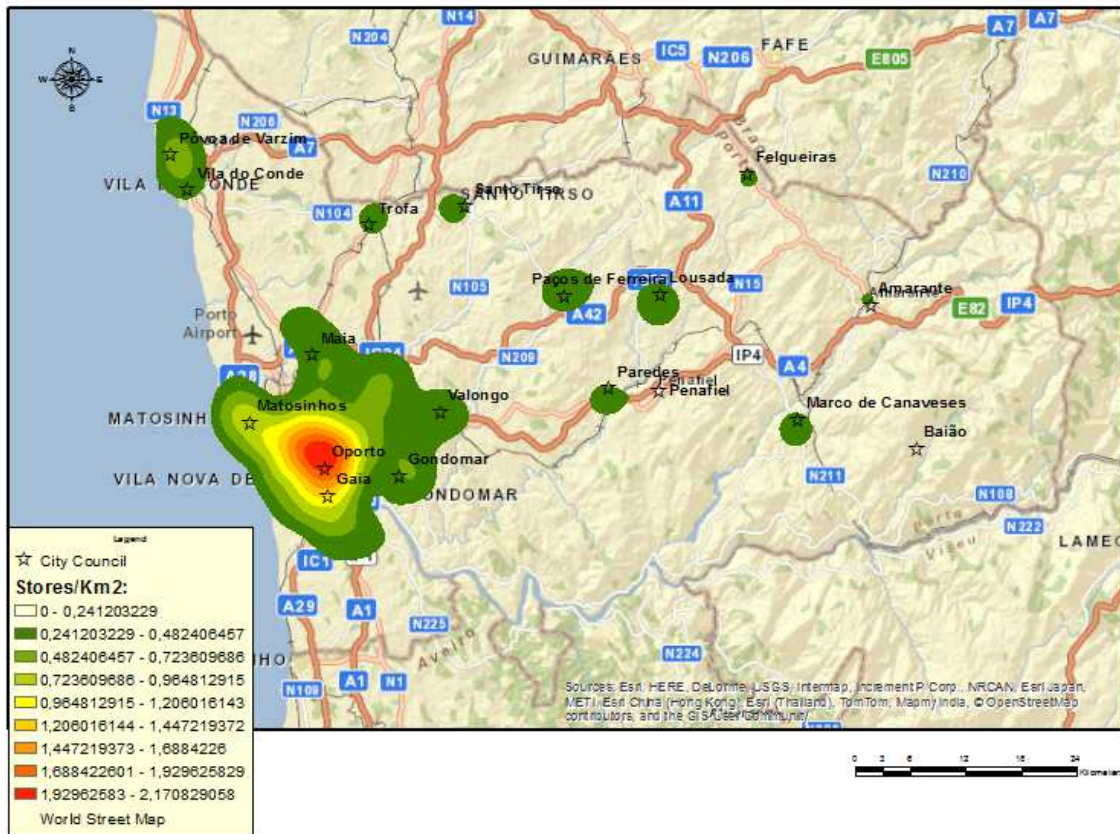
summarizes the results, with municipalities ordered according to their z-scores from the most to the least significant.

Table 5 - Average Nearest Neighbor Ratio of Oporto District Food Retailers

	Area (Km²)	Observed Mean Distance (meters)	Expected Mean Distance (meters)	Nearest Neighbor Ratio	z-score
Oporto District	2,331.7	788.87	1,461,25	0.54	-14.545**
<i>Municipality</i>					
Marco de Canaveses (7 stores)	201.89	330.01	2,685.21	0.123	-4.439**
Felgueiras (7 stores)	115.74	343.12	2,033.1	0.169	-4.207**
Trofa (5 stores)	71.88	288.72	1,895.75	0.152	-3.626**
Valongo (15 stores)	75.12	578.345	1,118.96	0.517	-3.579**
Vila do Conde (6 stores)	149.03	701.43	2,491.88	0.281	-3.367**
Amarante (5 stores)	301.33	1,041.1	3,881.57	0.268	-3.13*
Lousada (8 stores)	96.08	911.66	1,732.77	0.526	-2.564*
Baião (2 stores)	174.53	770.44	4,670.76	0.165	-2.259
Paços de Ferreira (9 stores)	70.99	981.96	1,404.29	0.699	-1.726
Matosinhos (27 stores)	62.42	662.94	760.24	0.872	-1.272
Santo Tirso (8 stores)	136.6	1,610.76	2,066.098	0.779	-1.192
Penafiel (6 stores)	212.24	2,431.54	2,973.79	0.818	-0.854
Gaia (44 stores)	168.47	1,034.29	978.36	1.057	0.725
Paredes (12 stores)	156.76	1,994.79	1,807.14	1.103	0.688
Póvoa de Varzim (10 stores)	82.21	1,532.695	1,433.58	1.069	0.418
Oporto (62 stores)	41.42	414.87	408.677	1.015	0.228
Gondomar (17 stores)	131.86	1,419.86	1,392.54	1.019	0.155
Maia (23 stores)	83.13	960.09	950.6	1.01	0.092
* Significant at 0.01					
** Significant at 0.001					

The kernel density estimation (KDE) method is used to calculate the density of features in a neighbourhood around those features (Silverman 1986), which in this case are the number of stores per square kilometre.

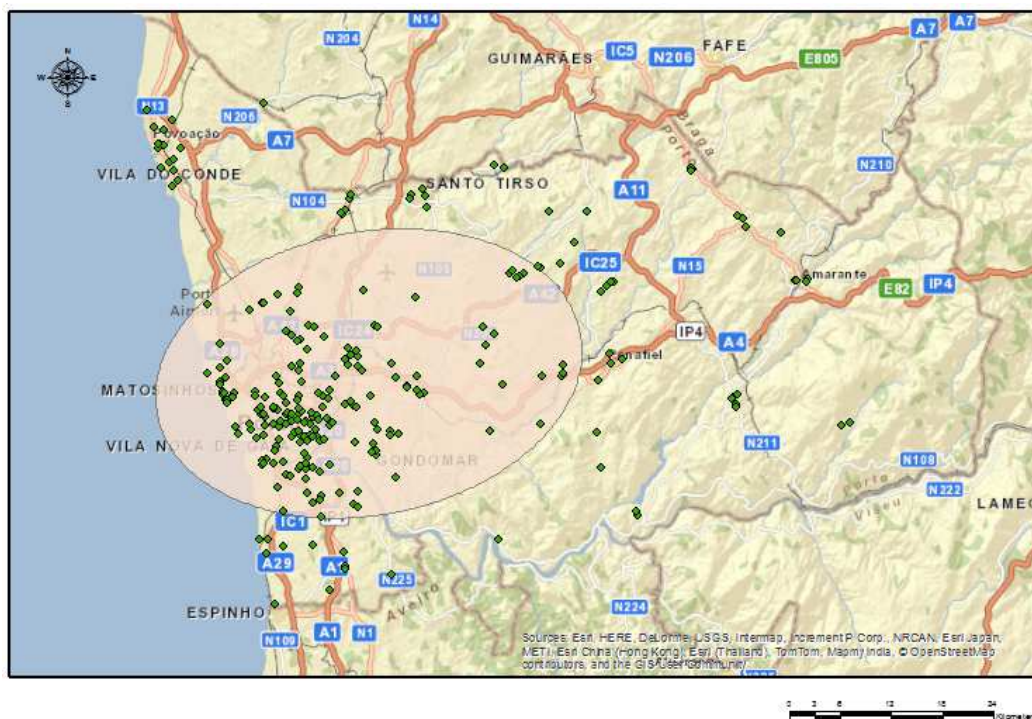
Figure 2 - Kernel Density Estimation Values of Oporto District Food Retailers



As shown in Figure 2 there's a clear higher density of stores around the most populated city (Oporto) where it's possible to find up to 2 stores per Km². With much less population, the other major cities in Oporto district have also a much lower density of stores. The KDE allow us to observe both centrality and agglomeration of stores in Oporto district since there's a concentration of stores around almost every city council, which lends support for both Central Place Theory and Principle of Minimum Differentiation.

The standard deviation ellipse measures the standard deviation of the distances of all stores from the mean centre thus measuring the spatial spread of their distribution (Mitchell 2005). The ellipse allows us to see if the distribution of stores is elongated and hence has a particular orientation. It is clear that the orientation of retail stores in Oporto district (Figure 3) is toward the larger city, Oporto, along the NE-SW direction, establishing it as the centre at the top of the hierarchy.

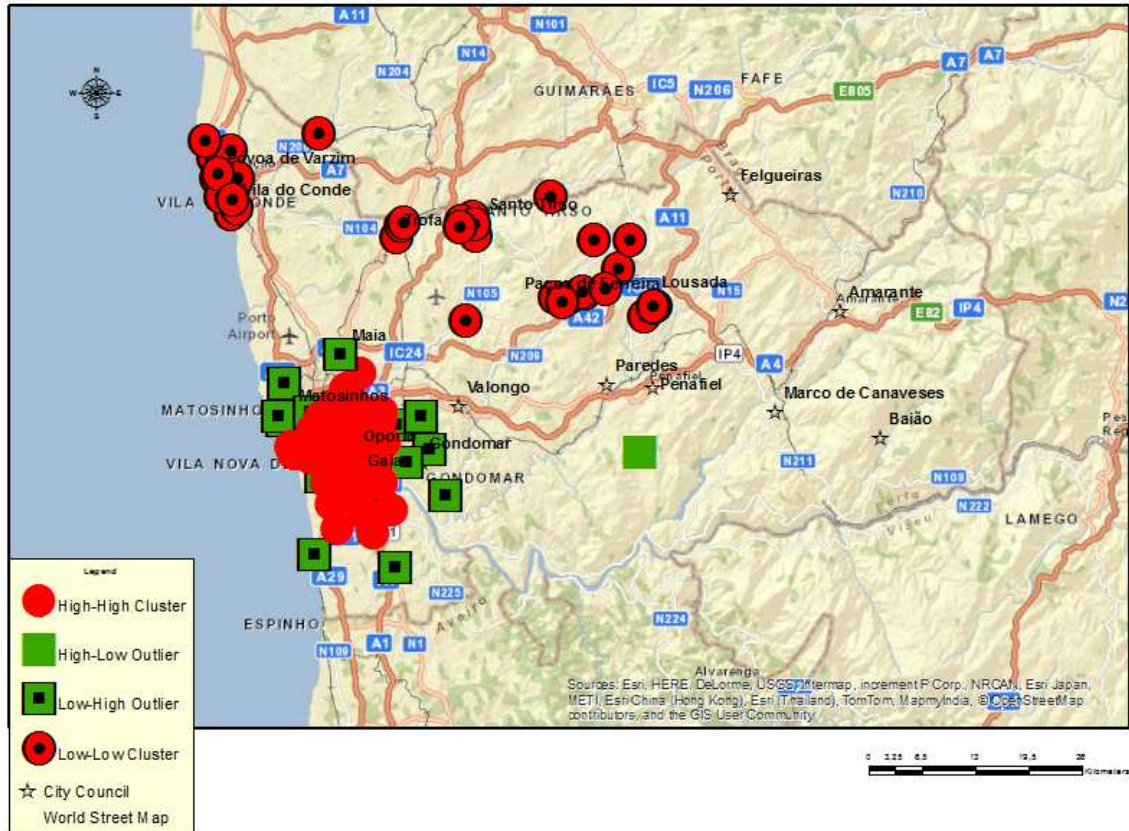
Figure 3 - Standard Deviation Ellipse of Oporto District Food Retailers



ArcGIS tool Cluster and Outlier Analysis identifies statistically significant hot spots, cold spots, and spatial outliers using the Anselin Local Moran's I statistic (Anselin 1995). A high positive z-score for a feature indicates that the surrounding features have similar values (either high values or low values). The COType will be HH for a statistically significant (0.05 level) cluster of high values and LL for a statistically significant (0.05 level) cluster of low values. A low negative z-score for a feature indicates a statistically significant (0.05 level) spatial outlier. The COType will indicate

if the feature has a high value and is surrounded by features with low values (HL) or if the feature has a low value and is surrounded by features with high values (LH). The high-high and low-low locations are typically referred to as spatial clusters, while the high-low and low-high locations are termed spatial outliers. The cluster is classified as such when the value at a location (either high or low) is more similar to its neighbours than would be the case under spatial randomness. As shown in Figure 4, several statistically significant clusters were identified, especially on the locations where the density of stores is higher indicating a tendency for agglomeration. Also the clusters tend to locate near the centre. The majority of outliers are recent stores. This could mean they are the first to arrive at new attractive locations and that in the future they could be part of a cluster.

Figure 4- Cluster and Outlier Analysis (Anselin Local Moran's I) of Oporto District Food Retailers



Finally, comparing the actual retailers spatial distribution to what would be expected from a random spatial distribution of the 273 stores will help to show if centrality and agglomeration of food retailers in Oporto district are strategic in nature or a result of randomness. Randomness is achieved by the Create Random Points tool from ArcGIS (see Appendix A). Create Random Points tool randomly places a specified number of points within an extent window or inside the features of a polygon, line, or point feature class (Esri 2013a). In this case the tool was used in each municipality of Oporto district to create the same number of random points as the number of existing stores. Four random sample of 273 stores were created. The average distance from the stores to the respective municipality city council, i.e. the centre, is 2,894 meters in the actual sample of stores and 5,105, 4,968, 4,955 and 5,012 meters in the random samples, indicating that, most likely, the proximity to the centre in the actual stores does not result from randomness. The same tests of density, directional distribution and cluster/outliers analysis were applied to the random samples.

Figure 5 - Kernel Density Estimation Values of Random Sample 1

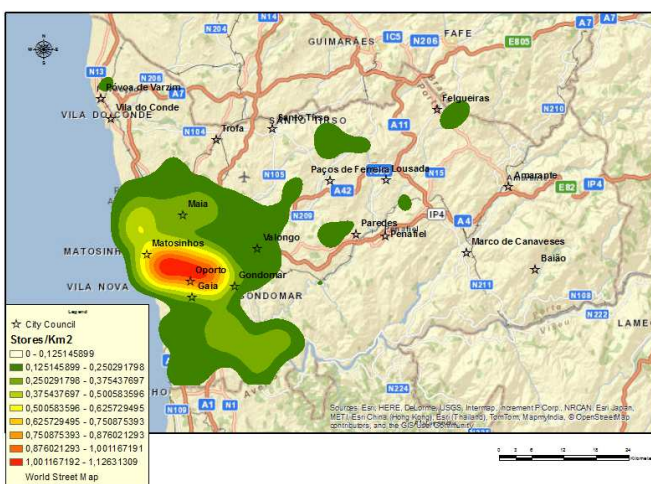


Figure 6 - Standard Deviational Ellipse of Random Sample 1

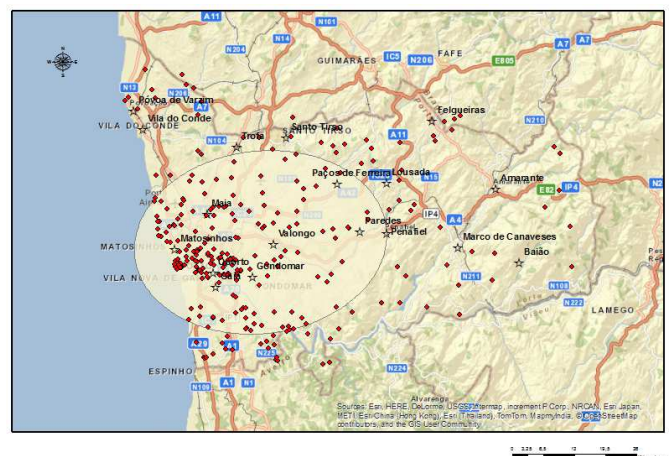
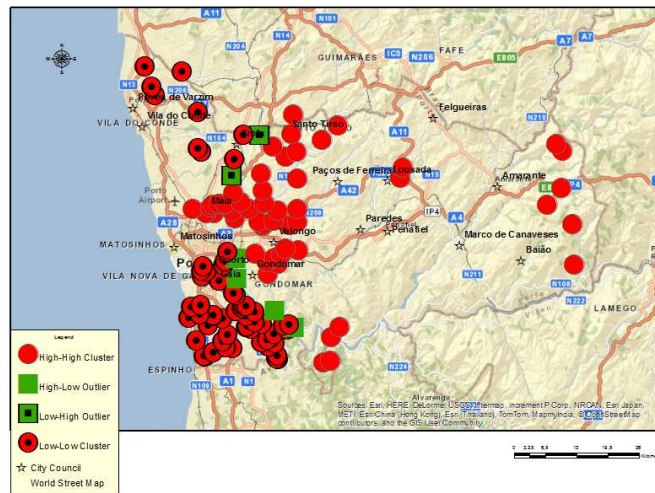


Figure 7 - Cluster and Outlier Analysis (Anselin Local Moran's I) of Random Sample 1



The results show a lower density of stores overall and a lower density and concentration around the centre (Figure 5), a weaker tendency towards the larger city with the stores being more scattered (Figure 6) and less concentration of clusters near the centre (Figure 7). The same tests on the other random samples resulted in very similar results (see Appendix B).

6. Conclusions

This exploratory research analyses the location pattern of food retailers (i.e. hypermarkets, supermarkets and discount stores) in Oporto district, Portugal, to test the predictive validity of three classic location theories: central place theory, spatial interaction theory and the principle of minimum differentiation, using several GIS techniques and statistical analysis. Given the preliminary nature of the study the results are not conclusive but the findings reported here do support the classic theories of retail location in the Portuguese case.

As predicted by Central Place Theory, store location is related to the store's location's surrounding population density. In our research this was verified by both correlation and

linear regression results that confirmed that the larger the population density the smaller the store's distance from the centre, as expected (Hypothesis 1a). The Kernel Density Estimation values also confirm the concentration of stores around almost every city council, where the population density is higher. The Cluster and Outlier Analysis (Anselin Local Moran's I) also confirms the clustering of stores around the centre. The proximity of stores to the centre does not result from randomness, as shown by the results of the Create Random Points tool. These results reinforce that in convenience shopping, like grocery shopping, proximity is important for the customer and that the retailers are aware of this. Central Place Theory also predicts that the larger the centre, namely in population, the larger the market area and the store density. The city of Oporto, being the most populated one, is at the top of the hierarchy with a clear store distributional trend towards it, as shown by the results of the Standard Deviational Ellipse. The city of Oporto has also the higher density of stores (2 stores per Km²), as shown by the KDE results. The correlation and linear regression results also confirmed the expected positive relationship between store density and population density (Hypothesis 1b) showing that, indeed, store location is related to the size of the centre.

As predicted by the Principle of Minimum Differentiation, store location is related to the competitor's store location. In our research we verified this premise through the correlation and linear regression results that confirmed that the larger the stores density the smaller the distance between competitors (Hypothesis 2a). The Nearest Neighbor Index also confirmed the clustering effect, in the whole district and also in each municipality, showing us also that there is a very low probability that the clustered pattern could be the result of random chance. These results also lend strength to Nelson's Theory of Cumulative Attraction since all the stores in our sample sell the same type of merchandise. As stated by Nelson (1958, p. 58), a given number of stores

dealing with the same merchandise will do more business if they are located adjacent, or in proximity to each other, than if they are widely scattered. The Principle of Minimum Differentiation also predicts that firms, when clustered, tend to locate near the centre of the market area. Once again, our results of both correlation and linear regression show the positive linear relationship between stores' distance to the nearest competitor and store's distance to the centre (Hypothesis 2b). Both store density and agglomeration are higher near the centre of each municipality, as shown by the KDE and Cluster and Outlier Analysis results.

Finally, as predicted in Spatial Interaction Theory, store location is related to consumer's proximity to the store (smaller stores are located in high population density locations) or the store's attractiveness to the consumer (larger stores, with more power of attraction, are located in lower population density locations). The correlation and linear regression results, when extreme values were removed (12 hypermarkets that are much larger than the majority of stores in the sample) confirm the expected negative relationship between store's size and population density. Where mobility or spatial elasticity of the customer is low, the gravitation of the customer tends to stores close to where they live or work. Where time is less important or mobility is higher, customers tend to gravitate towards larger stores located far from the centre.

If retailers are to adopt a more strategic approach to location it is crucial that they understand the micro and macro scale perspective of location spatial patterns. What we can conclude from this research is that the main location theories, from more than 30 years ago, are still relevant and can provide retailers key information to help optimize their location decisions.

The research looked only into food retailers, leaving out other types of retail formats. This option was mainly due to the fact that very few studies analyse this retail format and also because is the retail format with more stores thus providing a larger sample. It is however our intention to extend the research to other retail formats in the future. Another limitation of the present study relates to the use of the population density variable. This option was based primarily on the availability of this information from the 2011 Census and the assumption that the majority of customers derives from the store primary trading area (Dunne and Lusch 2007; Kotler, Armstrong, Saunders and Wong 2002; Levy and Weitz 2009). Although it is possible that one or more of the stores sampled could be patronized by significant number of customers from outside the immediate influence area, that possibility does not seem very likely since customers, when engaged in convenience shopping situations, are primarily concerned with minimizing their effort to get the product they want. In addition, there are serious limitations to the collection of customer information due to confidentiality issues or the absence of such information inside the company. A final limitation concerns the centre definition used. Despite the polycentric nature of Portuguese cities (Fernandes, Cachinho and Ribeiro 2000), city hall location continues to be the reference of city centre and is located inside or adjacent to the Central Business District. However other centre definitions can be used in future research such as the geographic centre or a spatial network structure to analyse centrality.

Acknowledgements

The authors acknowledge all the support given by the ESRI Portugal technical support team.

Appendix A. Process of Creating Random Numbers in ArcGIS

The first step of processing is to create a random number stream from a random number generator and seed. To randomly place the specified number of points in each polygon (each municipality), the polygons are partitioned by triangles of varying sizes using a standard polygon partitioning algorithm. The entire area of each polygon is filled by the triangles. To place the first point in the first polygon, one of the triangles in the overall polygon is randomly selected. Two legs of the triangle become the two axes from which to place the random point. The next unused value is taken from the random stream, transformed into the Uniform distribution using the start of the triangle axis as the minimum and the end of the leg as the maximum. A value is randomly selected on the axis. The same is done for the other axis or leg of the triangle. These two random values are used to place a point. The point will fall within a parallelogram created by the two axes of the triangle. This process is repeated until the specified number of points is placed in the polygon and repeated for each polygon. (Esri 2013b)

Appendix B. Figures with the results of tests performed to random samples

Figure 8 - Kernel Density Estimation Values of Random Sample 2

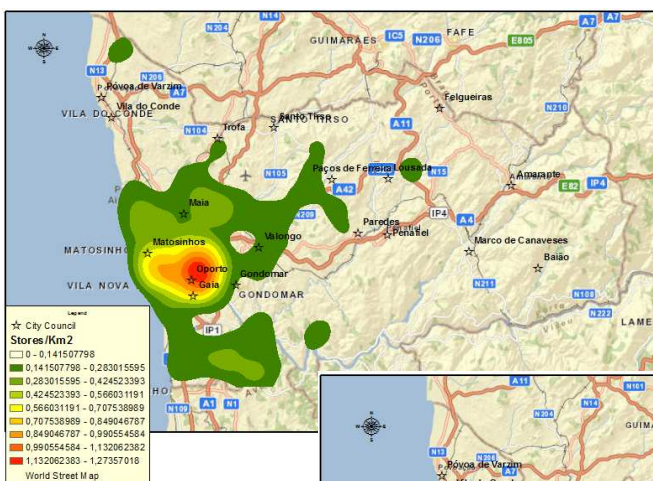


Figure 9 - Standard Deviation Ellipse of Random Sample 2

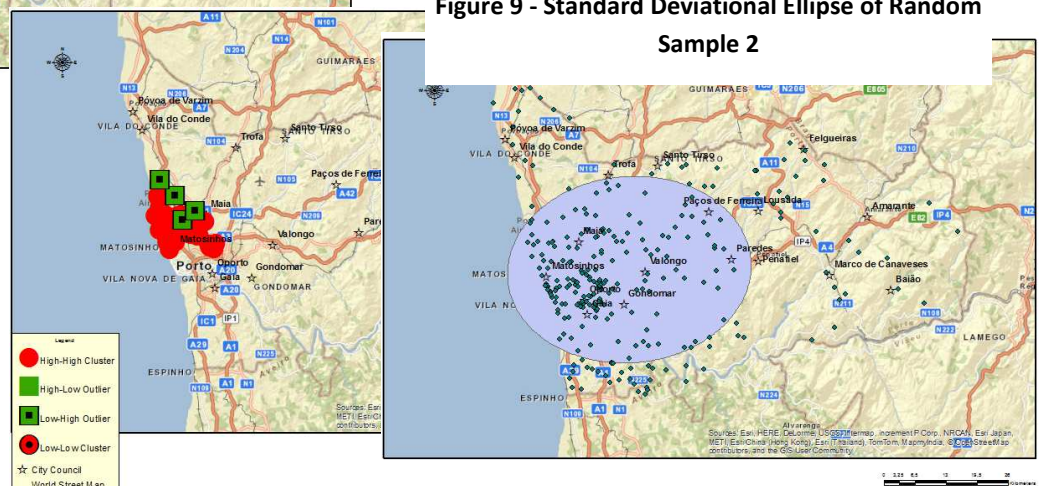


Figure 11 - Kernel Density Estimation Values of Random Sample 3

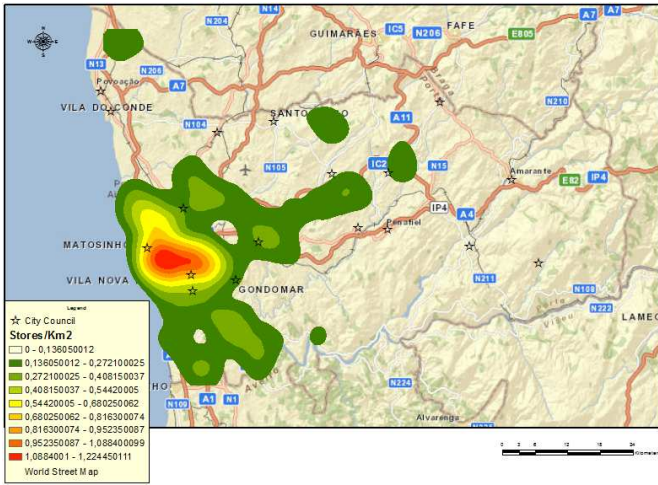


Figure 12 - Standard Deviation Ellipse of Random Sample 3

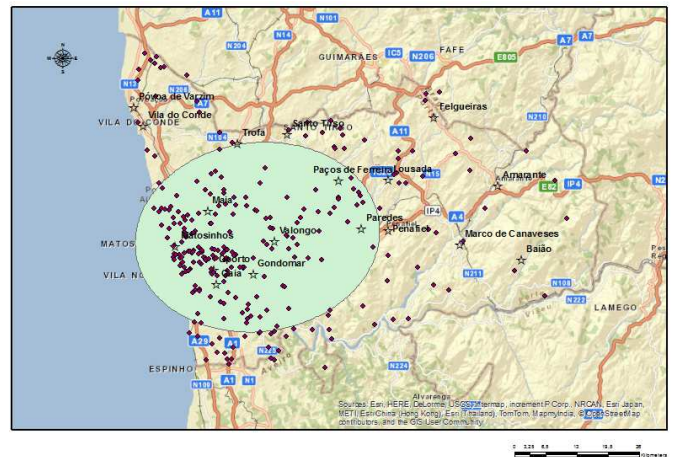


Figure 13 - Cluster and Outlier Analysis (Anselin Local Moran's I) of Random Sample 3

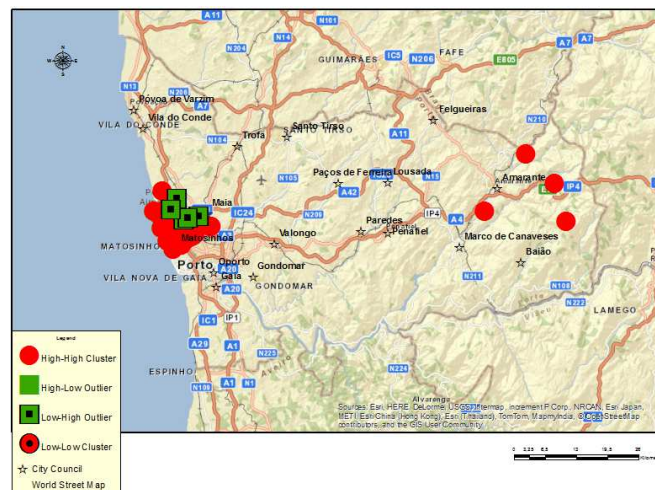


Figure 14 - Kernel Density Estimation Values of Random Sample 4

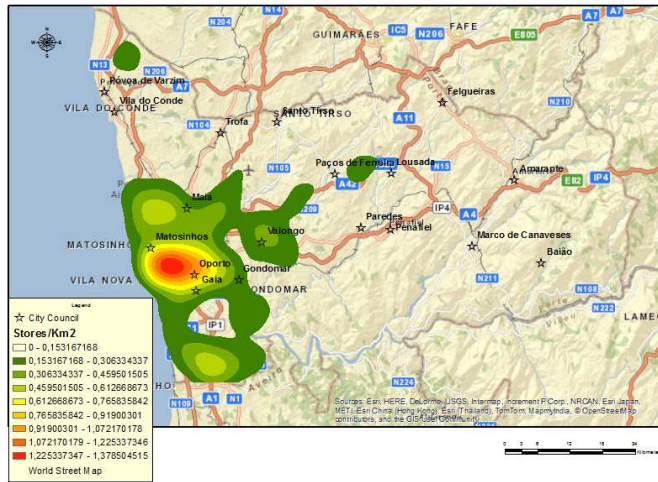


Figure 15 - Standard Deviation Ellipse of Random Sample 4

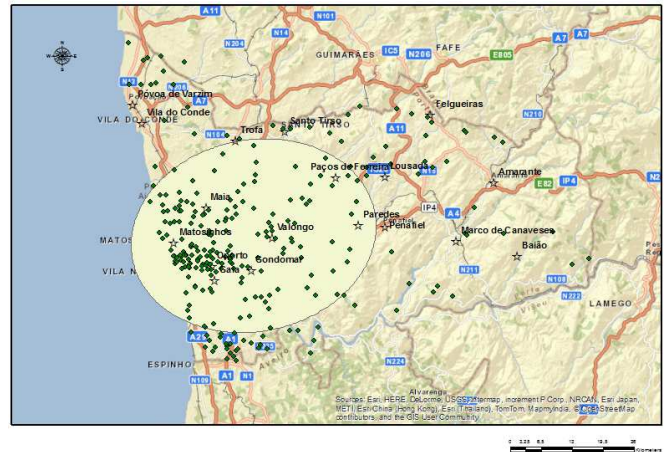
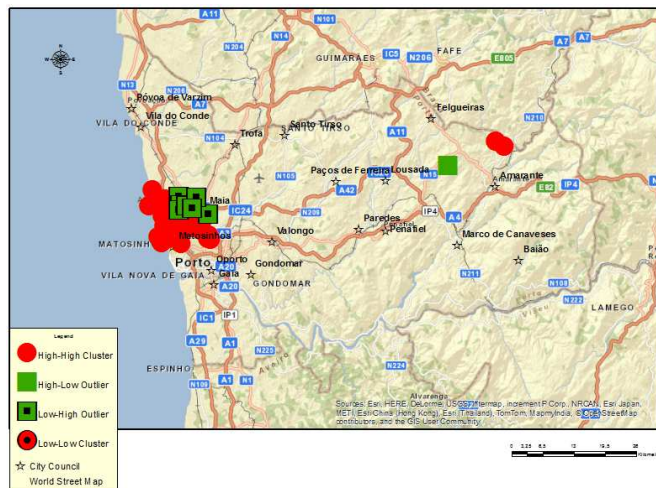


Figure 16 - Cluster and Outlier Analysis (Anselin Local Moran's I) of Random Sample 4



7. References

- Anselin, L. (1995). Local Indicators of Spatial Association - LISA. *Geographical Analysis*, 27(2), 93-115. doi: 10.1111/j.1538-4632.1995.tb00338.x
- Barreda-Tarrazona, I., García-Gallego, A., Georgantzís, N., Andaluz-Funcia, J., & Gil-Sanz, A. (2011). An experiment on spatial competition with endogenous pricing. *International Journal of Industrial Organization*, 29(1), 74-83. doi: 10.1016/j.ijindorg.2010.02.001
- Benoit, D., & Clarke, G. P. (1997). Assessing GIS for Retail Location Planning. *Journal of Retailing and Consumer Services*, 4(4), 239-258. doi: 10.1016/S0969-6989(96)00047-1
- Berry, B. J. L. (1963). Commercial structure and commercial blight: Retail patterns and progresses in the City of Chicago *Research Paper*, nr. 85: Department of Geography, University of Chicago.
- Beule, M. D., Poel, D. V. d., & Weghe, N. V. d. (2014). An extended Huff-model for robustly benchmarking and predicting retail network performance. *Applied Geography*, 46, 80-89. doi: 10.1016/j.apgeog.2013.09.026
- Birkin, M., Clarke, G., & Clarke, M. P. (2002). *Retail Geography and Intelligent Network Planning*. Chichester: Wiley.
- Borrell, J.-R., & Fernández-Villadangos, L. (2010). *Clustering or Scattering: the underlying reason for regulating distance among retail outlets*. Universitat de Barcelona. Xarxa de Referència en Economia Aplicada (XREAP). Não está completa. XREAP working paper n. 2010-12
- Brown, S. (1989). Retail Location Theory : The Legacy Of Harold Hotelling. *Journal of Retailing*, 65, 450-470.
- Brown, S. (1993). Retail Location Theory: Evolution and Evaluation. *The International Review of Retail, Distribution and Consumer Research*, 3, 185-229. doi: 10.1080/09593969300000014
- Brown, S. (1994). Retail Location at the Micro-Scale: Inventory and Prospect. *The Service Industries Journal*, 14, 542-576. doi: 10.1080/02642069400000056
- Cheng, E. W. L., Li, H., & Yu, L. (2007). A GIS approach to shopping mall location selection. *Building and Environment*, 42, 884-892.
- Christaller, W. (1933). *Central Places in Southern Germany*, translated by C. Baskin, 1966. Englewood Cliffs: Prentice Hall.
- Cox, R., & Brittain, P. (2004). *Retailing: An Introduction* (5ª ed ed.). Harlow: Pearson Education Limited.
- Cushman&Wakefield. (2012). *Marketbeat Research - Marketbeat Portugal*: Cushman & Wakefield. não está completa, pois há mais do que um relatório destes por ano.
- Cutsinger, J., Galster, G., Wolman, H., Hanson, R., & Towns, D. (2005). Verifying the Multi-Dimensional Nature of Metropolitan Land Use: Advancing the Understanding and Measurement of Sprawl. *Journal of Urban Affairs*, 27(3), 235-259. doi: 10.1111/j.0735-2166.2005.00235.x
- Daniels, M. J. (2007). Central Place Theory and Sport Tourism Impacts. *Annals of Tourism Research*, 34(2), 332-347. doi: 10.1016/j.annals.2006.09.004
- Davies, C. S., Holz, R. K., & Robertus, D. (1999). A Test of Central Place Theory using Shuttle Imaging Radar (SIR-A) of China's North Central Plain. *Geocarto International*, 14(1). doi: 10.1080/10106049908542089
- Davies, G., & Harris, K. (1990). *Small Business: The Independent Retailer*: Palgrave Macmillan.

- Dennis, C., Marsland, D., & Cockett, T. (2002). Central place practice: shopping centre attractiveness measures, hinterland boundaries and the UK retail hierarchy. *Journal of Retailing and Consumer Services*, 9, 185-199. doi: 10.1016/S0969-6989(01)00021-2
- Duggal, N. (2007). *Retail Location Analysis: A Case Study of Burger King & McDonald's in Portage & Summit Counties, Ohio.* não está completa Kent State University, College of Arts and Sciences, Department of Geography
- Dunne, P., & Lusch, R. (2007). *Retailing* (6th ed.): South-Western College Pub.
- Ebdon, D. (1986). *Statistics in Geography 2e: A Practical Approach* (2nd ed.): John Wiley & Sons.
- Esri. (2013a). Create Random Points (Data Management) Retrieved 15 March, 2015, from resources.arcgis.com/en/help/main/10.1/index.html#/Create_Random_Points/001700002r00000000/
- Esri. (2013b). How Create Random Points works Retrieved 15 March, 2015, from resources.arcgis.com/en/help/main/10.1/index.html#/How_Create_Random_Points_works/0017000000t4000000/
- Fernandes, J., Cachinho, H., & Ribeiro, C. (2000). *Comércio Tradicional em Contexto Urbano: Dinâmica de Modernização e Políticas Públicas.* Observatório do Comércio: Gabinete de Estudos para o Desenvolvimento e Ordenamento do Território - Faculdade de Letras da Universidade do Porto.
- Galster, G., Hanson, R., Ratcliffe, M. R., Wolman, H., Coleman, S., & Freihage, J. (2001). Wrestling Sprawl to the Ground: Defining and Measuring an Elusive Concept. *Housing Policy Debate*, 12(4), 681-717. doi: 10.1080/10511482.2001.9521426
- Ghosh, A. (1994). *Retail Management* (2nd ed.): Dryden Press.
- González-Benito, Ó. (2005). Spatial competitive interaction of retail store formats: modeling proposal and empirical results. *Journal of Business Research*, 58, 457-466. doi: 10.1016/j.jbusres.2003.09.001
- Haig, R. M. (1927). *Regional Survey of New York and its Environs.* Nova Iorque: New York City Planning Commission.
- Hausman, J., & Leibtag, E. (2007). Consumer benefits from increased competition in shopping outlets: Measuring the effect of Wal-Mart. *Journal of Applied Econometrics*, 22(7), 1157-1177. doi: 10.1002/jae.994
- Hotelling, H. (1929). Stability in Competition. *The Economic Journal*, 39, 41-57.
- Hsu, W.-T. (2012). Central Place Theory and City Size Distribution. *The Economic Journal*, 122(563), 903-932. doi: 10.1111/j.1468-0297.2012.02518.x
- Huang, A., & Levinson, D. (2008). *An agent-based retail location model on a supply chain network.* University of Minnesota: Nexus Research Group. não seria melhor substituir este relatório pelo artigo? Why retailers cluster? An agent model of location choice.
- Hymel, K. (2014). Do parking fees affect retail sales? Evidence from Starbucks. *Economics of Transportation*, 3, 221-233. doi: 10.1016/j.ecotra.2014.08.001
- INE. (2013a). *Censos* 2011. http://censos.ine.pt/xportal/xmain?xpid=CENSOS&xpgid=censos2011_apresentacao
- INE. (2013b). *Contas Nacionais Trimestrais.* In C. Q. Detalhados) (Ed.).
- Irmen, A., & Thisse, J.-F. (1998). Competition in Multi-Characteristics Spaces: Hotelling was almost right. *Journal of Economic Theory*, 78, 76-102. doi: 10.1006/jeth.1997.2348
- Kotler, P., Armstrong, G., Saunders, J., & Wong, V. (2002). *Principles of Marketing* (3rd ed.): Pearson Education Limited.
- Lee, M.-L., & Pace, R. K. (2005). Spatial Distribution of Retail Sales. *The Journal of Real Estate Finance and Economics*, 31, 53-69. doi: 10.1007/s11146-005-0993-5
- Levy, M., & Weitz, B. (2009). *Retailing Management* (7th ed.). Nova Iorque: McGraw-Hill/Irwin.

- Li, Y., & Liu, L. (2012). Assessing the impact of retail location on store performance: A comparison of Wal-Mart and Kmart Stores in Cincinnati. *Applied Geography*, 32, 591-600. doi: 10.1016/j.apgeog.2011.07.006
- Litz, R. a., & Rajaguru, G. (2008). Does Small Store Location Matter? A Test of Three Classic Theories of Retail Location. *Journal of Small Business & Entrepreneurship*, 21, 477-492. doi: 10.1080/08276331.2008.10593436
- Mafra, F., & Silva, J. A. (2004). *Planeamento e Gestão do Território*. Porto. está incompleto, falta a editora: Sociedade Portuguesa de Inovação
- McGoldrick, P. (2002). *Retail Marketing*. Nova Iorque: McGraw-Hill Education.
- Mendes, A. B. (2005). *Modelação de Vendas de Novas Superfícies Comerciais (PhD Thesis)*. Incompleta: Universidade Técnica de Lisboa
- Mitchell, A. (2005). *The Esri Guide to GIS Analysis, Volume 2: Spatial Measurements and Statistics* (Vol. 2): Esri Press.
- Murad, A. A. (2011). Creating a GIS Application for Retail Facilities Planning in Jeddah City. *Journal of Computer Science*, 7(6), 902-908. doi: 10.1016/S0303-2434(03)00020-5
- Nakamura, D. (2014). Social participation and social capital with equity and efficiency: An approach from central-place theory. *Applied Geography*, 49, 54-57. doi: 10.1016/j.apgeog.2013.09.008
- Nelson, R. (1958). *The Selection of Retail Locations*. Nova Iorque: F.W. Dodge Corp.
- Netz, J. S., & Taylor, B. A. (2002). Maximum and Minimum Differentiation? Location Patterns of Retail Outlets. *The Review of Economics and Statistics*, 84(1), 162-175.
- Nogueira, M., Crocco, M., Figueiredo, A. T., & Diniz, G. (2014). Financial hierarchy and banking strategies: a regional analysis for the Brazilian case. *Cambridge Journal of Economics*, 39(1), 139-156. doi: 10.1093/cje/beu008
- Önden, I., Sen, C. G., & Sen, A. (2012). Integration of Integer Programming with GIS Analyzing Abilities for Determining the Convenience Levels of Retail Stores. *Procedia - Social and Behavioral Sciences*, 62, 1144-1149. doi: 10.1016/j.sbspro.2012.09.196
- Pence, J. (2011). *Measuring Omaha, Nebraska's Urban Sprawl from 1990-2000*: Proquest, UMI Dissertation Publishing.
- Pennerstorfer, D., & Weiss, C. (2013). Spatial Clustering and Market Power: Evidence from the retail gasoline market. *Regional Science and Urban Economics*, 43, 661-675. doi: : 10.1016/j.regsciurbeco.2013.04.002
- Picone, G. A., Ridley, D. B., & Zandbergen, P. A. (2009). Distance decreases with differentiation: Strategic agglomeration by retailers. *International Journal of Industrial Organization*, 27, 463-473. doi: 10.1016/j.ijindorg.2008.11.007
- Porta, S., Strano, E., Iacoviello, V., Messori, R., Latora, V., Cardillo, A., . . . Scellato, S. (2009). Street centrality and densities of retail and services in Bologna, Italy. *Environment and Planning B*, 36, 450-465. doi: 10.1068/b34098
- Reilly, W. J. (1929). *Methods for the Study of Retail Relationships*. Austin: University of Texas, Bureau of Business Research.
- Reilly, W. J. (1931). *The Law of Retail Gravitation*. Nova Iorque: W.J. Reilly.
- Reimers, V., & Clulow, V. (2004). Retail concentration: a comparison of spatial convenience in shopping strips and shopping centres. *Journal of Retailing and Consumer Services*, 11, 207-221. doi: 10.1016/S0969-6989(03)00038-9
- Rodrigues, D., Vala, F., & Monteiro, J. (2002). Áreas de Influência das Cidades de Lisboa e Vale do Tejo *Região de Lisboa e Vale do Tejo: Áreas de Influência dos Serviços e Sistema Urbano Regional*: Comissão de Coordenação Regional de Lisboa e Vale do Tejo.
- Roig-Tierno, N., Baviera-Puig, A., Buitrago-Vera, J., & Mas-Verdu, F. (2013). The retail site location decision process using GIS and the analytical hierarchy process. *Applied Geography*, 40, 191-198. doi: 10.1016/j.apgeog.2013.03.005
- Scott, P. (1959). The Australian CBD. *Economic Geography*, 35, 290-314.

- Scott, P. (1970). *Geography and Retailing*. Chicago: Aldine Pub. Co.
- Shields, M., & Kures, M. (2007). Black out of the blue light: An analysis of Kmart store closing decisions. *Journal of Retailing and Consumer Services*, 14(259-268). doi: 10.1016/j.jretconser.2006.07.007
- Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. New York: Chapman & Hall.
- Suárez-Vega, R., Santos-Peñate, D. R., & Dorta-González, P. (2012). Location models and GIS tools for retail site location. *Applied Geography*, 35, 12-22. doi: 10.1016/j.apgeog.2012.04.009
- Themido, I. H., Quintino, A., & Leitão, J. (1998). Modelling the Retail Sales of Gasoline in a Portuguese Metropolitan Area. *International Transactions in Operational Research*, 5(2), 89-102. doi: 10.1111/j.1475-3995.1998.tb00106.x
- Turhan, G., Akalin, M., & Zehir, C. (2013). Literature Review on Selection Criteria of Store Location Based on Performance Measures. *Procedia - Social and Behavioral Sciences*, 99, 391-402. doi: 10.1016/j.sbspro.2013.10.507
- Wang, F., Chen, C., Xiu, C., & Zhang, P. (2014). Location analysis of retail stores in Changchun, China: A street centrality perspective. *Cities*, 41, 54-63. doi: 10.1016/j.cities.2014.05.005
- Wolman, H., Galster, G., Hanson, R., Ratcliffe, M., Furdell, K., & Sarzynski, A. (2005). The Fundamental Challenge in Measuring Sprawl: Which Land Should Be Considered? *The Professional Geographer*, 57(1), 94-105. doi: 10.1111/j.0033-0124.2005.00462.x
- Zhang, Z. J. (1995). Price-Matching Policy and the Principle of Minimum Differentiation. *The Journal of Industrial Economics*, 43(3), 287-299. doi: 10.2307/2950581