

Um Estudo dos Supermercados no Brasil: Uma Investigação sobre a Área de Influência

Autoria: Juracy Parente, Heitor Takashi Kato

RESUMO

O trabalho procura dar uma contribuição ao conhecimento varejista no Brasil, integrando aspectos teóricos e empíricos de importantes temas ligado ao varejo: localização varejista, e Área de Influência. Procurou-se inicialmente, a partir da revisão da literatura, desenvolver uma conceituação sobre a Área de Influência e sua importância para o setor de supermercados. Esse estudo está baseado em extensa pesquisa com dados coletados junto a mais de 20 mil clientes de uma amostra de 27 supermercados localizados em São Paulo. Um processo de geo-processamento, utilizando o software MapInfo, permitiu calcular a distância entre loja e a residência de cada consumidor. O trabalho apresenta os resultados que descrevem e mensuram a Área de Influência dos supermercados pesquisados. Uma análise de "cluster" conseguiu identificar a existência de 3 agrupamentos de curvas representantes do padrão de área de influência das 27 lojas da amostra. Na parte final, foi desenvolvido um modelo econométrico que permite explicar e prever as dimensões e características da Área de Influência das lojas. Os resultados indicam que o modelo tem um alto grau de poder explicativo das Áreas de Influência, permitindo prever com alto grau de precisão estas áreas.

INTRODUÇÃO E OBJETIVOS

Os conceitos de gestão adotados nas empresas varejistas têm sido, quase em sua maioria, elaborados nos países desenvolvidos, especialmente nos Estados Unidos, onde se produz a maior parte do conhecimento e da literatura sobre o assunto. No Brasil, a produção do conhecimento nessa área não tem sido muito grande. Observa-se uma carência de trabalhos que possam ajudar a estabelecer uma ligação mais profunda entre os conceitos e os arcabouços teóricos com a realidade do varejo brasileiro. Essa escassez de publicações no Brasil dificulta ao varejista brasileiro a utilização desses conceitos na tomada de decisões empresariais.

A motivação para o desenvolvimento dessa pesquisa consiste exatamente em dar uma contribuição para o preenchimento dessa lacuna. Esse trabalho procura dar continuidade e aprofundar pesquisas anteriores, objetivando desenvolver uma investigação de temas de importância prioritária para o varejista brasileiro: localização varejista e Área de Influência.

Os seguintes fatores contribuíram para a escolha desse tema: a) grande importância que os varejistas nacionais dedicam a esse assunto; b) carência de pesquisas realizadas sobre o tema no Brasil e, conseqüentemente, pouco material didático desenvolvido para os varejistas brasileiros.

Reconhecendo que os conceitos são as unidades fundamentais das teorias, pode-se considerar que este trabalho visa auxiliar no desenvolvimento de conceitos relacionados com o fenômeno do desempenho varejista. Ainda em 1964, Kaplan (1964, p.53-54) indica que conceitos adequados são necessários para a formulação de uma boa teoria. Adicionalmente, uma boa teoria é necessária para o desenvolvimento de bons conceitos. Kaplan denomina de “paradoxo da conceitualização” essa situação de interdependência entre o desenvolvimento de conceitos e o desenvolvimento de teoria. Ele sugere que esse paradoxo vá sendo resolvido por um círculo virtuoso de aproximação, ou seja, à medida que conceitos venham sendo mais bem definidos, a teoria vai avançando, e, progressivamente, com o avanço da teoria, os conceitos vão sendo melhor estabelecidos. Esse projeto visa exatamente contribuir para o avanço desse círculo virtuoso de aproximação e ajudar a superar esse “paradoxo de conceitualização”, que vem caracterizando a ainda embrionária teoria varejista.

Os objetivos específicos desse trabalho são:

- Descrever a Área de Influência de 27 supermercados de São Paulo, ou seja o fenômeno da dispersão geográfica de clientes em torno da loja;
- Operacionalizar por meio de indicadores numéricos e de curvas o conceito da Área de Influência.
- Procurar encontrar generalizações no fenômeno da distribuição geográfica dos clientes em torno das lojas
- Desenvolver modelos quantitativos que ajudem a explicar e prever os raios de Área de Influência.

Esse artigo está estruturado em 6 partes. Inicialmente apresentamos a revisão do conhecimento de Área de Influência, e em seguida mostramos a metodologia de pesquisa utilizada para a coleta dos dados empíricos. A terceira parte está dedicada para apresentar os resultados da dispersão geográfica dos clientes nos 27 supermercados pesquisados, e na operacionalização do conceito de área de influência. Na quarta parte, apresentamos os resultados das análises de “cluster” e analisamos as regularidades e generalizações encontradas no fenômeno da área de influência. Em seguida mostramos a estrutura e os resultados do modelo de regressão múltipla desenvolvido para explicar e prever a área de influência. E, finalmente, a última parte é dedicada às conclusões do trabalho e para as sugestões para novos estudos.

1- REVISÃO DO CONHECIMENTO

Um dos fatores chaves para o sucesso de qualquer empresa varejista consiste na escolha adequada da localização de suas lojas. A literatura varejista vem desenvolvendo um considerável esforço no sentido de entender os vários aspectos que determinam a atratividade dos pontos varejistas. Os pesquisadores na área de localização varejistas vêm dedicando uma especial atenção ao conceito da Área de Influência e às metodologias para seu mapeamento e análise.

A Área de Influência é um conceito fundamental nos estudos de localização, pois reflete a dimensão espacial da demanda de mercado do varejo. Costumamos defini-la como sendo a “área geográfica que contém a maior parte dos clientes de uma loja”. Segundo a American Marketing Association (AMA), a Área de Influência é uma área geográfica contendo os consumidores de uma empresa particular ou grupo de empresas para bens ou serviços específicos¹. Esse conceito tem sido pesquisado há bastante tempo com os estudos de Reilly(1929), McGill (1933) e de Converse(1949) e começou a ser estudado mais intensamente a partir da década de 60, com os trabalhos apoiados na analogia de Applebaum (1966) e nos modelos gravitacionais de Huff(1964). A análise da dispersão geográfica dos clientes em torno de uma loja permitiu identificar 3 segmentos de uma Área de Influência (Applebaum, 1966):

- Primária – cerca de 60 a 75% dos clientes que moram mais perto da loja;
- Secundária – cerca de 15 a 25% dos clientes que moram em torno da área primária;
- Terciária – parcela restante, cerca de 10% dos clientes que moram mais afastados da loja.

Os modelos gravitacionais deram grande contribuição para o conhecimento sobre a Área de Influência. Esses modelos estão baseados nas leis de física de gravitação dos corpos celestes, desenvolvidas por Isaac Newton, ainda no século XVII. Huff (1964) desenvolveu um modelo que não fica restrito a um único contorno delimitador de influência, mas define diferentes contornos para diferentes probabilidades do cliente comprar em certa loja. A formulação de

Huff, portanto, consiste em uma adaptação mais realista do modelo gravitacional, já que considera simultaneamente as múltiplas alternativas de compra que o consumidor seleciona.

De acordo com essa formulação, a probabilidade de que as compras ocorressem em um determinado ponto de venda a partir de um dado ponto dependeria diretamente do poder de atração deste ponto de venda e seria inversamente proporcional à distância da moradia (ou ponto de partida) em relação àquele. O trabalho de Huff foi importante, pois foi o primeiro que sugeriu que as áreas de influência eram complexas, contínuas e probabilísticas², definidas por meio de contornos que apresentavam uma mesma probabilidade (90% , 60%, 5%) de consumidores a serem atraídos para a loja.

Os estudos sobre Áreas de Influência vêm também ganhando renovado interesse com o desenvolvimento da metodologia do *Geographical Information System* (GIS)(Birkim, Clark, Clark e Wilson, 1996). A metodologia do GIS permite integrar um grande volume de informações censitárias (perfil demográfico e econômico por micro-regiões) aos mapas geográficos computadorizados, facilitando o desenvolvimento de análises de mercado e do mapeamento das áreas de influência das lojas que operam em certo mercado. Neste estudo, faremos uso da metodologia GIS para mapear a Área de Influência dos 27 supermercados pesquisados.

2 - METODOLOGIA DE PESQUISA

Neste estudo, adotaremos uma metodologia com uma abordagem quantitativa mais efetiva, já que um de nossos objetivos consiste em procurar desenvolver modelos econométricos que possam explicar e prever as características e dimensões da Área de Influência. Os detalhes metodológicos de nossa investigação são explicados a seguir.

ÁREA DE INFLUÊNCIA E *CUSTOMER SPOTTING*

A investigação sobre a Área de Influência foi conduzida por meio de uma pesquisa empírica com consumidores de supermercados de uma rede varejista brasileira, utilizando a metodologia do *customer spotting*, ou seja, o mapeamento da residência de clientes em torno de cada loja pesquisada.

A amostra aleatória dos clientes foi selecionada utilizando o cadastro de clientes. O mapeamento foi efetuado pela localização do endereço da residência de cada um, com a técnica de “georeferenciamento”. Mapas computadorizados (versão “MapInfo”) da Grande São Paulo foram utilizados para indicar no mapa a localização da residência de cada cliente.

POPULAÇÃO E AMOSTRA DE LOJAS

A população de varejistas é constituída pelos supermercados na Grande São Paulo.

Dois motivos foram decisivos na escolha dos supermercados para realizar esse projeto de pesquisa:

- Enorme importância do modelo supermercado no varejo brasileiro;
- Experiência prévia e facilidade do pesquisador em ter acesso a empresas supermercadistas.

A cidade de São Paulo foi selecionada devido aos seguintes fatores:

- Grande importância dessa região no mercado brasileiro;
- Disponibilidade de mapas computadorizados e de dados econômico-demográficos dos distritos e de micro-setores da região;
- Proximidade do pesquisador e, conseqüentemente, a facilidade de realizar pesquisas de observação nas lojas e regiões pesquisadas.

As lojas pesquisadas refletem uma amostra representativa de supermercados de São Paulo, já que englobam unidades com diferentes características, em termos de: tamanho de loja, densidade populacional, nível socioeconômico e grau de concorrência. Foram

selecionados 27 supermercados localizados na cidade de São Paulo, representando quase a totalidade de lojas que a empresa pesquisada opera na região. Esse tamanho de amostra foi considerado satisfatório para o desenvolvimento dos modelos econométricos. O Quadro 3 apresenta algumas características dessas 27 lojas, tais como tamanho (m² da área de vendas) e densidade populacional das áreas no entorno das lojas.

AMOSTRA DE CLIENTES PARA CADA LOJA

A investigação da Área de Influência foi realizada por meio do mapeamento da residência dos clientes em torno de cada loja. O tamanho de amostra apresentou uma média de 900 clientes por loja, possibilitando resultados com margens de erro confortáveis.

Com base na formulação de tamanho de amostra³,

$$S_p = \sqrt{\frac{pxq}{n}}$$

estimamos que, com 95% de segurança, as margens de erro para resultados percentuais (ilustrando: % de clientes localizados dentro do raio de 2 km) oscila entre 2 e 3%.

Para este estudo, utilizou-se o software estatístico SPSS®, V. 10.01.

3 – RESULTADOS DE MENSURAÇÕES E OPERACIONALIZAÇÃO DO CONCEITO DE ÁREA DE INFLUÊNCIA

ÁREA DE INFLUÊNCIA SUMARIZADA POR MEIO DE TABELAS E INDICADORES

O Quadro 1, com raios predefinidos traçados em torno da loja, mostra a percentagem acumulada de clientes dentro de cada um desses raios. Essa alternativa proporciona uma valiosa informação agregada, refletindo o padrão geral da distribuição geográfica em torno da loja.

É interessante destacar como as diversas lojas apresentam diferente capacidade de atrair clientes. Existem lojas com áreas de influência muito compactas, como as lojas 1, 2, 9, e 10, que concentram mais de 60% de seus clientes dentro do raio de 600 metros, enquanto outras lojas apresentam áreas de influência bem ampliadas, como as lojas 12, 20 e 27, cujos raios (que contém 60% de seus clientes) se estendem até os 3 km.

Assim, 60% parece ser um percentual representativo da Área de Influência, pois inclui a maioria dos clientes, sem, entretanto, incorporar áreas com uma densidade daqueles muito baixa. Com o objetivo de procurar sumarizar o fenômeno da Área de Influência por meio de um único indicador, definiremos o indicador de Área de Influência como sendo o raio que contém 60% dos clientes mais próximos em torno da loja.

QUADRO 1 - PORCENTAGEM ACUMULADA DE CLIENTES EM RAIOS CONCÊNTRICOS PREDEFINIDOS (FONTE: PESQUISA)

Código Loja	Percentagens acumuladas de clientes								
	10%	20%	30%	40%	50%	60%	70%	80%	90%
1	90	142	190	250	302	403	538	760	2.181
2	81	154	249	334	423	517	670	998	5.262
3	237	333	463	717	2.098	4.408	8.751	12.122	17.350
4	322	423	568	830	1.149	1.993	3.166	4.775	9.454
5	214	321	429	553	721	929	1.108	1.702	3.798
6	373	527	780	1.065	1.318	1.629	2.150	3.011	4.366
7	204	315	434	547	662	833	1.161	2.255	5.257
8	142	306	453	598	760	939	1.368	3.726	9.902
9	104	196	250	276	393	529	740	1.196	3.396
10	139	187	255	308	404	425	489	696	6.898
11	601	729	898	1.158	1.744	2.299	3.043	3.946	6.670
12	199	332	485	773	1.087	1.491	3.163	5.670	9.629
13	205	338	452	565	767	1.116	1.452	2.350	5.705
14	172	253	325	394	453	598	1.186	5.615	12.206
15	256	433	588	799	960	1.656	5.975	11.506	16.070
16	375	541	712	891	1.102	1.340	1.746	2.814	4.345
17	334	550	781	1.060	1.359	2.250	4.139	7.440	11.361
18	836	1.082	1.269	1.471	1.655	1.949	2.776	3.691	6.672
19	133	235	436	621	1.066	3.218	6.476	9.983	12.450
20	504	800	1.053	1.441	1.656	1.967	2.440	3.655	5.591
21	0	153	277	359	458	581	774	1.062	2.514
22	247	486	1.099	2.078	4.613	5.904	7.284	9.078	10.891
23	257	359	438	515	583	686	771	873	1.009
24	244	408	546	703	983	1.435	2.260	4.959	10.603
25	207	375	547	701	893	1.122	1.661	3.943	8.675
26	302	529	708	897	1.030	1.252	1.631	2.210	3.886
27	525	847	1.255	1.791	2.247	2.727	3.315	4.638	7.582

ÁREA DE INFLUÊNCIA SUMARIZADA POR MEIO DE CURVAS

Com base na natureza cumulativa dos dados desenvolvemos uma representação gráfica para proporcionar um sumário mais completo dos fenômenos da distribuição dos clientes em torno de lojas. Ao capturar a maior parte das informações contidas nos mapas de *customer spotting*, os gráficos ilustram uma forma útil de representar o padrão de concentração e dispersão dos clientes. O eixo horizontal representa a distância em linha reta entre a loja e a residência do cliente, e o eixo vertical representa a percentagem acumulada do cliente dentro de cada intervalo de distância.

Esse tipo de diagrama representa um excelente e completo resumo do fenômeno da dispersão geográfica dos clientes, pois, em qualquer ponto da curva, uma relação específica é estabelecida entre o raio da distância (entre loja e residência de cliente) e a percentagem de clientes contida nesses raios. A natureza cumulativa do eixo vertical ajuda a oferecer respostas a questões práticas perguntadas por varejistas, tais como: De qual distância uma percentagem de clientes (60%, por exemplo) vem? Qual percentagem de clientes vem de uma certa distância (raio de 2 km, por exemplo)? Essas curvas oferecem uma visão precisa da relação espacial entre a loja e seus clientes, e reflete a capacidade da primeira em atrair os segundos.

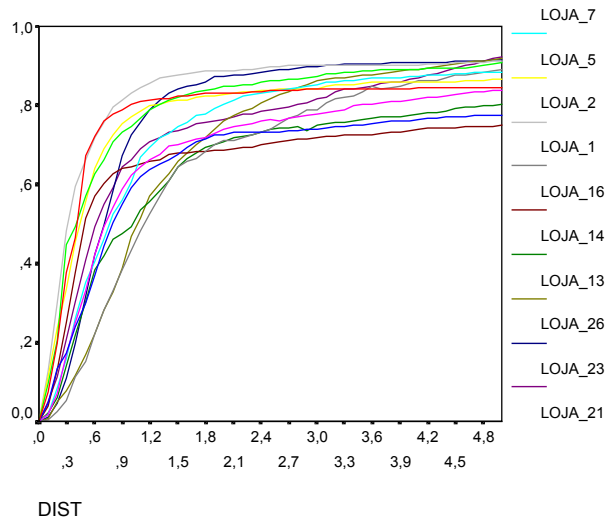
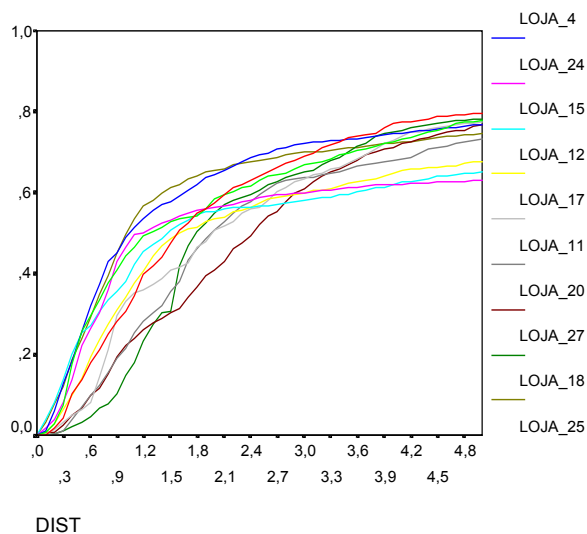


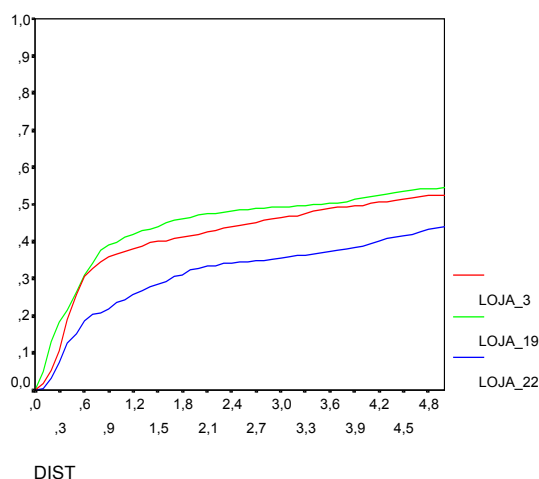
FIGURA 2 - CURVAS DE ÁREA DE INFLUÊNCIA (FONTE: PESQUISA)

	<p>Lojas do Grupo 1 –</p> <p>As 13 lojas do Grupo 1 (Vide Figura 2) têm Área de Influência mais concentrada e compacta, típica de supermercados de porte pequeno, localizados em regiões de alta densidade populacional. 60% dos clientes estão contidos no raio de 400 metros a 1 km. A curva sobe em alta inclinação até o nível do 70%, e apresenta uma sensível mudança de inclinação entre 70 e 80%, quando assume uma característica assintótica</p>
--	---



Lojas do Grupo 2 –

As curvas das 11 lojas do Grupo 2 (Figura 3) mostram um padrão claramente menos concentrado, quando comparado às das lojas do Grupo 1. Apresentam uma inclinação mais suave, mesmo nos segmentos iniciais da curva, e começam a mostrar sua natureza assintótica entre 60 a 70%. Em geral, são lojas, localizadas em regiões de níveis médios de densidade populacional, cujos raios precisam se estender até 2 km para manter a maioria dos clientes.



	<p>Lojas do Grupo 3 -</p> <p>As lojas do Grupo 3 (Figura 4) apresentam uma curva muito alongada e refletem características peculiares. São supermercados pequenos, localizados em regiões com concentração comercial (lojas e/ou escritórios), próximos a terminais ou corredores de transporte coletivo. As curvas desse grupo de lojas refletem um público heterogêneo, constituído por dois segmentos de clientes.</p>
--	--

Para identificar a existência de agrupamentos de lojas, caracterizados por padrões semelhantes de comportamento em relação às curvas acumuladas de clientes, procedeu-se uma análise de *cluster* (análise de agrupamentos) das lojas, com base na proporção acumulada de clientes a cada 100 metros para tentar avaliar o comportamento conjunto das mesmas. Usamos as distâncias das lojas como variáveis características e procedemos a uma análise do dendrograma resultante. A análise sugere que a maioria das lojas pode ser classificada em 2 categorias principais. Um terceiro grupo de 3 lojas com comportamento atípico foi também identificado. A figura 1 apresenta o dendrograma resultante da análise de “*cluster*”.

GENERALIZAÇÕES SOBRE FENÔMENO DA ÁREA DE INFLUÊNCIA

A inspeção dessas curvas mostra que cada loja apresenta uma relação curvilínea diferente, refletindo seu comportamento particular de atrair clientes. Apesar dessas variações, algumas regularidades podem ser observadas na maioria das lojas, nos Grupos 1 e 2:

1. A inclinação da curva, apesar de alta nos segmentos mais próximos da loja, declina rapidamente além de um certo ponto – 50 a 70%. De fato, podemos identificar, de acordo com a inclinação das curvas, 3 segmentos:
 - a. um segmento com inclinação alta – de 50% a 70%;
 - b. um segmento transicional em que a inclinação acusa uma sensível modificação, caindo rapidamente – entre 60 e 80%;
 - c. além da oscilação de 70 a 80%, quando a inclinação da curva torna-se muito pequena, assumindo uma característica assintótica.

2. Essa distinção de segmentos é mais nítida nas curvas das lojas do Grupo 1, em geral, lojas menores em regiões mais densas. Nas lojas do Grupo 2, geralmente, lojas maiores localizadas em regiões menos densas, a identificação dos três segmentos não é muito clara.
3. A natureza assintótica, entretanto, foi encontrada em todas as curvas – o que mostra a existência habitual de pequena percentagem de clientes que são atraídos devido à conveniência da loja em relação aos seus deslocamentos, referentes aos seus locais de trabalho ou pela proximidade dos terminais de transporte
4. A observação desse padrão de dispersão oferece uma confirmação empírica: a classificação da Área de Influência nos 3 segmentos, conforme proposto por Applebaum⁴ há quase 40 anos, ainda prevalece atualmente, pois a descrição do comportamento da maioria das curvas (Grupo 1 e Grupo 2) de Área de Influência coincide muito de perto com os três segmentos que foram identificados na relação curvilínea de nossa pesquisa.

5 - MODELO PREDITIVO DO RAIOS DA ÁREA DE INFLUÊNCIA

A estrutura de qualquer modelo deve ser definida por meio de hipóteses sobre as relações que existem entre as variáveis. Os vários autores que estudaram e desenvolveram modelos sobre Área de Influência identificaram muitos fatores significativos no padrão e na dimensão da Área de Influência na estrutura que estabelece as relações entre as variáveis “dependentes” e as variáveis independentes.

- Mix Varejista da Loja;
 - o Tamanho da Loja;

Com base na pesquisa bibliográfica e no conhecimento advindo das pesquisas empíricas já realizadas, as seguintes variáveis foram selecionadas para explicar as dimensões da Área de Influência:

- Variáveis Ambientais;
 - o Densidade Populacional;
 - o Presença de Centros Comerciais e Terminais de Transporte Coletivo.

TAMANHO DA LOJA – “TAM”

Os modelos⁵ comentados anteriormente utilizam o tamanho da loja como variável *proxi* para refletir a atratividade da unidade varejista. O modelo gravitacional está baseado na hipótese de que a atratividade de um centro varejista varia proporcionalmente com a variedade de produtos oferecida por esse centro. Na maioria das formulações baseadas em modelos gravitacionais, o tamanho da loja, refletindo a sua variedade e sortimento, é a variável utilizada para operacionalizar a atratividade da loja. Efetivamente, lojas maiores apresentam, em geral, uma variedade mais expressiva, mais conforto e serviços, e preços muito competitivos, obtendo, assim, uma capacidade mais elevada para atrair clientes.

Rosembloom (1976) afirma que um aumento na variedade deverá ampliar a distância que consumidores estariam dispostos a percorrer para alcançar uma certa loja. O autor sugere que essa influência começa a ser decrescente a partir de um certo nível de variedade, além do qual um acréscimo em variedade não conseguirá exercer uma atratividade mais efetiva nos clientes. Esse mesmo ponto de vista é defendido por outros autores⁶, que afirmam que, embora a extensão da Área de Influência seja afetada pelo tamanho da loja, sua extensão não cresce proporcionalmente ao aumento de tamanho daquela, a relação é não-linear.

Em nosso estudo, o tamanho da loja foi medido por meio de sua área de vendas (m²). A informação sobre o tamanho da loja foi obtida com a rede de supermercados que participou do projeto.

DENSIDADE POPULACIONAL – “DENS”

É grande a influência que a densidade populacional exerce no desempenho de uma loja. A densidade populacional afeta a demanda de mercado disponível e, portanto, determina o número de clientes que freqüentam uma certa loja, sua Área de Influência e o seu volume de vendas.

Mas, além do seu impacto na demanda de mercado, a densidade populacional também afeta o padrão de comportamento dos consumidores em seu processo de compra em supermercados. Parece razoável conjecturar que uma densidade populacional mais alta tenderá a inibir a distância que os consumidores estariam dispostos a percorrer para fazerem suas compras. A percepção da distância é afetada pela densidade, já que o tráfego tende a estar mais pesado e andar em velocidade mais lenta em regiões mais congestionadas. Um outro efeito da densidade é que regiões mais povoadas atraem uma oferta varejista mais elevada e, portanto, os consumidores encontram um leque de alternativas mais perto de casa. Essas conjecturas são confirmadas em bairros verticalizados, como Higienópolis em São Paulo, ou Copacabana no Rio de Janeiro, onde os consumidores já se habituaram a reduzir o uso de seus veículos para irem as compras, e muitos preferem andar para alcançarem as lojas próximas.

Neste estudo, a densidade populacional foi determinada no raio de 1 km em torno da loja, e foi estimada com dados extraídos do último censo (2001) do IBGE, e reflete o número de habitantes por hectare⁷, localizado dentro do raio de 1 km em torno da loja.

TRANSPORTE COLETIVO – “COLET”

A análise Figura 2 mostra que algumas lojas (lojas 3, 19 22) apresentam um comportamento atípico, com áreas de influência muito estendidas, refletindo características peculiares. São lojas localizadas em regiões com concentração comercial ou de escritórios, próximas a estações ou corredores de transporte coletivo. Ficou evidente, portanto, que esse fator deveria ser incorporado ao nosso modelo. Criou-se, então, uma variável *dummy* para refletir essa outra variável, adotando-se valor 1 para a presença de terminais ou grande disponibilidade de transporte coletivo, e o valor zero para regiões sem essa característica.

MODELOS DE REGRESSÃO

Para a estimação dos parâmetros do modelo preditivo, utilizou-se o modelo de regressão linear múltipla, em que a formulação geral toma a seguinte forma⁸:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + \varepsilon_i, \text{ onde}$$

Y= variável dependente

β_1 = constante

β_k = coeficientes das variáveis independentes

X_k = variáveis independentes

ε = termo de erro

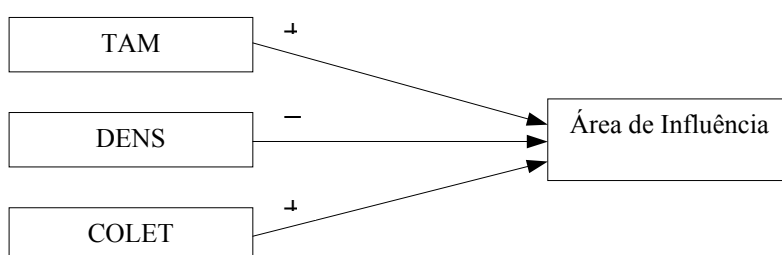
Para este estudo, utilizou-se o software estatístico SPSS®, V. 10.01.

A variável dependente é a dimensão da Área de Influência (“AI-60”). Conforme explicado e justificado, essa variável será operacionalizada por meio do comprimento do raio (metros) que contém 60% da residência dos clientes da loja.

As variáveis usadas nessa estimação foram as variáveis “DENS”, “TAM” e “COLET”, conforme explicado anteriormente. O modelo proposto está ilustrado na Figura 5, e as hipóteses que estão norteando o desenvolvimento desse modelo são:

- H₁: A extensão da Área de Influência (“AI”) está positivamente relacionada com o Tamanho da Loja (“TAM”)
- H₂: A extensão da Área de Influência (“AI”) está negativamente relacionada com a Densidade Populacional no entorno da loja (“DENS”).
- H₃: A extensão da Área de Influência (“AI”) está positivamente relacionada com a presença de concentração comercial e transporte coletivo (“COLET”)

FIGURA 3 - MODELO PREDITIVO DE ÁREA DE INFLUÊNCIA



RESULTADOS DA REGRESSÃO

O Quadro 2, apresenta os resultados do Modelo 1, alternativa selecionada para se estimar Área de Influência. O Modelo 1 foi calibrado usando-se todas as 27 lojas pesquisadas. Após analisar os resultados de diferentes alternativas, o melhor modelo encontrado expressa Log de Área de Influência como função das três variáveis: TAM, LogDENS e COLET. O Quadro 2, abaixo, traz os resultados da regressão, apresentando os coeficientes e os valores do teste *t* de Student.

QUADRO 2 - RESULTADOS DO MODELO 1 (FONTE: PESQUISA)

	Const	TAM	LogDENS	COLET	R ² Ajust	F	n
Coef	4,30	1,29E-4	0,70	0,46	0,735	24,98	27
t	11,69	3,12	-4,26	5,64			
Sig	0,000	0,005	0,00	0,000		0,000	

O Modelo 1 mostra resultados muito satisfatórios, pois apresenta um bom Coeficiente de Correlação Ajustado de 0.735 (indicando que 73,5% da variação na extensão da Área de Influência pode ser explicado por variações no conjunto de variáveis independentes), e também com elevadíssimo valor F (24,98), que denota um auto grau de ajuste do modelo. Verifica-se também que os sinais (positivo ou negativo) dos coeficientes de todas as variáveis são consistentes com as direções indicadas nas nossas expectativas prévias (sinalizadas por meio das hipóteses), e são também significantes – com seus respectivos valores do teste *t* de Student, acima do nível de 99% de significância. As três hipóteses propostas sobre os fatores que afetam a Área de Influência ficam, assim, confirmadas.

ANÁLISE DOS RESÍDUOS

A análise dos resíduos do Modelo 1, apresentado no Quadro 3, revela uma loja com resíduo maior do que 2 desvios padrão. O modelo subestimou a Área de Influência da Loja 3

(resíduo = - 2,44 desvios padrão). É uma loja pequena, situada próxima à área de grande densidade populacional. A região é também caracterizada por forte concentração de grandes prédios de escritórios e bancos, com expressiva população flutuante. A extensa Área de Influência da loja é provocada pela grande massa de população flutuante que trabalha na região, que parece utilizar a loja ao saírem de seus escritórios.

A loja com o 2º maior desvio é a Loja 4, também subestimada pelo modelo (resíduo = -1,7 desvio padrão). Trata-se de uma unidade localizada em um importante eixo viário que serve como principal via de interligação entre vários bairros da cidade. A grande avenida onde a loja está localizada fortalece a capacidade de atrair clientes de regiões mais distantes.

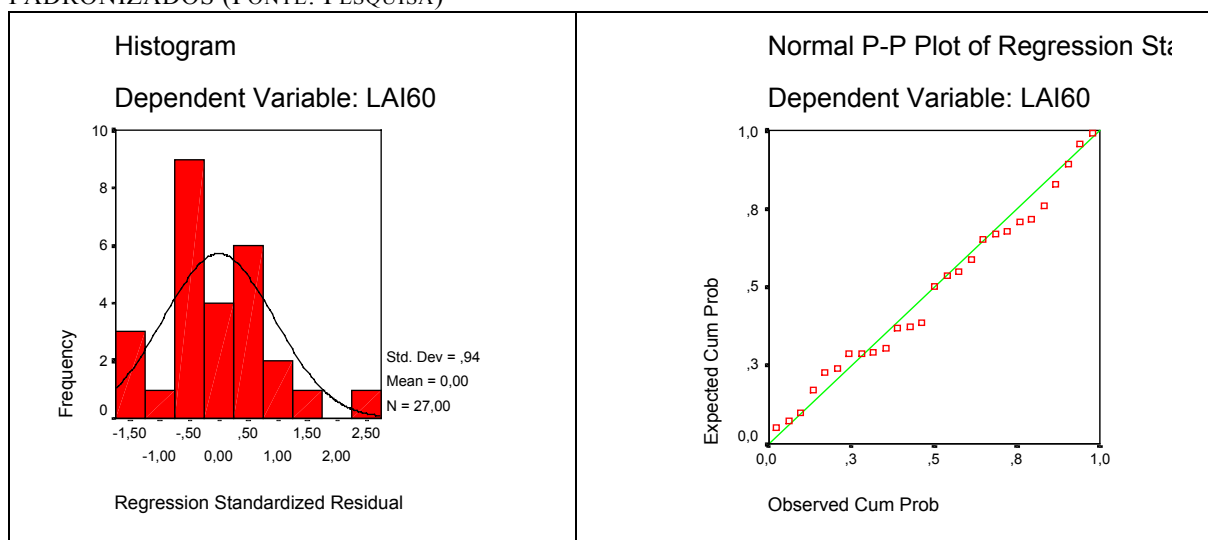
Identificamos ainda duas outras lojas com desvios expressivos, ambas com áreas de influência superestimadas pelo modelo. A Loja 12 (desvio = -1,6 desvios) apesar de estar situada em uma travessa próxima (50 metros) de um importante corredor de ônibus, fica muito escondida, e pouco visível para o intenso movimento de pedestres da avenida (corredor de ônibus). Existe também um supermercado concorrente de grande porte, localizado logo na esquina da avenida (corredor de ônibus), bem mais acessível e visível para o intenso movimento de pedestres que circula pelo corredor de ônibus. Esses dois fatores contribuíram para inibir o poder de atração da Loja 12. A Loja 1 (desvio = -1,5) está localizada em via secundária de uma região estritamente residencial e, assim, tem seu poder de atração restrito à população que reside em suas proximidades.

QUADRO 3 - MATRIZ COM DADOS DAS 27 LOJAS, PREVISÕES E RESÍDUOS DO MODELO
(FONTE: PESQUISA)

Nº Id Loja	AI-60	Tam M2	Dens Hab/ha	Colet (Dum-my)	Log Dens	Log AI-60	Prev AI-60	Res AI-60	Prev AI-60 Padroniz	Res AI-60 Padroniz
1	403	882	173	0	2,24	2,61	2,84	-0,23	-1,03	-1,46
2	517	442	263	0	2,42	2,71	2,65	0,06	-1,71	0,39
3	4.408	440	162	1	2,21	3,64	3,26	0,38	0,57	2,44
4	1.993	1.178	103	0	2,01	3,30	3,03	0,27	-0,29	1,72
5	929	1.536	113	0	2,05	2,97	3,05	-0,08	-0,23	-0,51
6	1.629	2.465	130	0	2,11	3,21	3,13	0,09	0,07	0,55
7	833	1.311	164	0	2,21	2,92	2,91	0,01	-0,76	0,09
8	939	1.127	130	0	2,11	2,97	2,95	0,02	-0,58	0,12
9	529	1.200	147	0	2,17	2,72	2,93	-0,20	-0,69	-1,29
10	425	554	202	0	2,31	2,63	2,75	-0,12	-1,36	-0,75
11	2.229	1.959	63	0	1,80	3,35	3,28	0,07	0,64	0,44
12	1.491	470	94	1	1,97	3,17	3,43	-0,26	1,21	-1,63
13	1.116	1.248	75	0	1,88	3,05	3,14	-0,09	0,10	-0,56
14	598	890	157	0	2,20	2,78	2,87	-0,09	-0,91	-0,57
15	1.656	743	179	1	2,25	3,22	3,27	-0,05	0,61	-0,33
16	1.340	1.700	126	0	2,10	3,13	3,04	0,09	-0,27	0,58
17	2.250	1.225	103	1	2,01	3,35	3,50	-0,15	1,47	-0,94
18	1.949	943	34	0	1,53	3,29	3,33	-0,04	0,85	-0,29
19	3.218	349	88	1	1,94	3,51	3,43	0,07	1,22	0,47
20	1.967	4.080	103	0	2,01	3,29	3,41	-0,11	1,11	-0,71
21	581	1.136	184	0	2,26	2,76	2,85	-0,09	-0,97	-0,54
22	5.904	1.050	39	1	1,59	3,77	3,77	0,00	20,48	0,00
23	686	1.202	167	0	2,22	2,84	2,89	-0,05	-0,83	-0,32
24	1.435	1.432	124	0	2,09	3,16	3,01	0,15	-0,38	0,95
25	1.122	596	143	0	2,16	3,05	2,86	0,19	-0,95	1,24
26	1.252	1.610	111	0	2,05	3,10	3,06	0,03	-0,17	0,22
27	2.727	3.186	92	0	1,96	3,44	3,32	0,11	0,81	0,71

Os gráficos dos resíduos padronizados do Modelo 1, apresentados a seguir, mostram características de distribuição normal, o que respeita os pressupostos do modelo, conforme demonstrado na Figura 4.

FIGURA 4 - HISTOGRAMA DOS RESÍDUOS PADRONIZADOS E GRÁFICO DOS RESÍDUOS PADRONIZADOS (FONTE: PESQUISA)



MULTICOLINEARIDADE

Como a precisão dos modelos de regressão pode tornar-se severamente comprometida pela existência de intercorrelação entre seu conjunto de preditores, precisamos analisar esses resultados para eliminar os casos de alta multicolinearidade. O Quadro 4 apresenta a matriz de correlação entre as variáveis do Modelo 1. Os resultados indicam que não existem casos de alta multicolinearidade, e portanto, esse conjunto de preditores poderá ser utilizado sem o risco de violar os pressupostos do modelo.

Quadro 4 - MATRIZ DE CORRELAÇÃO (FONTE: PESQUISA)

Correlations		LAI60	Tam-m2	LDENS	Colet
LAI60	Pearson Correlation	1	,208	-,659	,600
	Sig. (2-tailed)	-	,297	,000	,001
	N	27	27	27	27
Tam-m2	Pearson Correlation	,208	1	-,202	-,376
	Sig. (2-tailed)	,297	-	,313	,053
	N	27	27	27	27
LDENS	Pearson Correlation	-,659	-,202	1	-,196
	Sig. (2-tailed)	,000	,313	-	,328
	N	27	27	27	27
Colet	Pearson Correlation	,600	-,376	-,196	1
	Sig. (2-tailed)	,001	,053	,328	-
	N	27	27	27	27

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

6 - CONCLUSÕES

Reconhecendo que os conceitos são as unidades fundamentais das teorias, pode-se considerar que esse trabalho atingiu seu objetivo de auxiliar no desenvolvimento de constructos da ainda embrionária teoria varejista, visando uma melhor descrição, entendimento, explicação e previsão de indicadores dos fenômenos relacionados ao desempenho varejista, avançando na resolução do que Kaplan(1964) chama de “paradoxo da conceitualização”.

Os resultados da mensuração do fenômeno da distribuição geográfica dos clientes dos 27 supermercados demonstram que existe uma enorme diferença na capacidade que diferentes lojas tem de atrair clientes. Enquanto que para algumas lojas, 60% dos clientes estão localizados dentro de um raio de 500 metros, para outras unidades essa percentagem está contida em raios maiores que 3 km.

A análise de “cluster” permitiu a identificação de dois principais agrupamentos de curvas representantes do padrão de área de influência de 24 lojas, cerca de 90% ds lojas de nossa amostra. Algumas generalizações nas curvas puderam ser observadas em todos os agrupamentos das lojas. A análise das curvas identificou 3 segmentos: um segmento inicial, com uma inclinação mais alta, um segmento onde ocorre uma transição na inclinação da curva (entre 50% e 70% dos clientes) e um terceiro segmento, a partir do qual a curva assume uma natureza assintótica.

Uma das grandes vantagens constatadas de se reduzir a Área de Influência em indicadores quantitativos foi a de permitir o desenvolvimento de modelos quantitativos para explicar e prever Área de Influência como função de um conjunto de variáveis preditoras, que influenciam o comportamento e a preferência dos consumidores para supermercados.

Os resultados do modelo econométrico apresentam-se satisfatórios não só para explicar os fatores que determinam uma maior ou menor Área de Influência, mas também parecem adequados para prever as dimensões daquela. Os resultados nos levam a afirmar com alto grau de confiança que a Área de Influência é positivamente afetada pela variável TAM e COLET, e negativamente influenciada pela DENS. As estimativas apresentam uma grande coerência com a proposta dos modelos gravitacionais, os quais estão baseados na influência de dois fatores: um fator de atração e um fator inibidor. O fator de atração tem sido, geralmente, operacionalizado por meio da variável tamanho da loja. O fator inibidor é expresso pela distância entre a loja e a residência (ou ponto de partida) do cliente. A distância reflete o custo ou a dificuldade desse deslocamento. Dificuldades mais elevadas no deslocamento atuarão, por um lado, como inibidores da distância a ser percorrida. Por outro lado, a facilidade mais acessível no deslocamento incentivará o percurso de distâncias mais longas.

As variáveis “DENS” e “COLET” incorporadas ao modelo refletem o grau de dificuldade que a distância representa. A variável “DENS” representa a dificuldade de se chegar perto da loja com o uso de automóveis: quanto maior a densidade, maior o volume de tráfego e maior resistência para se chegar ao ponto de venda. Já a variável “COLET” reflete a facilidade de acesso à loja com o uso de meios de transporte coletivo: a existência próxima à loja de corredores de transporte coletivo, ônibus ou metrô, diminui a dificuldade de se chegar ao ponto de venda.

A análise dos resíduos revelou que nosso modelo ainda não consegue retratar adequadamente o comportamento e o impacto da “população flutuante” no desempenho dos supermercados. Essa situação ficou especialmente evidente naquelas lojas que sofrem a influência de grandes contingentes de população flutuante, pelo fato de estarem localizadas em pólos de comércio, ou em regiões de concentração de escritórios, em áreas próximas a terminais de transporte coletivos. Sugerimos que novas pesquisas sejam realizadas para uma

melhor compreensão da dimensão e do impacto desse segmento de consumidores no desempenho dos supermercados.

As regularidades encontradas no fenômeno da distribuição geográfica dos clientes em torno da loja indicam que novos estudos sejam realizados para investigar não só quais tipos de funções curvilíneas podem refletir o comportamento das curvas de área de influência, mas também o desenvolvimento de modelos matemáticos que possam explicar e prever o formato dessas curvas.

REFERÊNCIAS

- APPLEBAUM, W.. "Methods for determining store trade areas, market penetration and potential sales". *Journal of Marketing Research*, Vol. III, May, 1966, pág. 127-141.
- BELL, David R; HO, Teck-Hua; TANG, Christopher S. "Determining where to shop: Fixed and variable costs of shopping", *Journal of Marketing Research*, Aug 1998 p: 352-369.
- BENNETT, Peter D.. (org.). *Dictionary of Marketing Terms*, 2nd ed.. Chicago, American Marketing Association, 1995, pg. 287. In: BERMAN, Barry; EVANS, Joel R.. *Retail Management: A Strategic Approach*. Upper Saddle River, Prentice Hall, 1998, pg. 276.
- BERMAN, Barry; EVANS, Joel R.. *Retail Management: A strategic Approach*. Upper Saddle River, Prentice Hall, 1998.
- BIRKIN Mark, CLARK, Graham, CLARKE, Martin; WILSON Alan. *Intelligent GIS: Location Decision and Strategic Planning*. New York, John Wiley & Sons, 1996.
- CONVERSE, Paul D.. "New Laws of Retail Gravitation", *Journal of Marketing*, 14 (October), 1949, pg. 379-384.
- CRAIG, C. Samuel; GHOSH, Avijit; McLAFFERTY, Sara. "Models of the Retail Location Process: A Review". *Journal of Retailing*, Vol. 60, no.1, Spring 1984, pg. 5-36.
- HAIR, Joseph F., ANDERSON, Rolph E., TATHAM, Ronald L. BLACK, William C. *Multivariate data analysis*. New Jersey : Prentice-Hall, 1998.
- HUFF, David L.. "Defining and Estimating a Trade Area". *Journal of Marketing*, July 1964, pág.37.
- KAPLAN, A. , *The Conduct of Inquiry: Methodology for Behavioral Sciences*, New York: Intext Educational Publishers, Chandler Publishing Company, 1964.
- McGILL, Kenneth H.. "A Method for Delineating Retail Trade Area". *Journal of Retailing*, April 1933, pg. 10-14.
- PINDYCK, Robert S.; RUBINFELD, Daniel L.. *Econometric Models & Economic Forecasts*. 3rd. Ed.. New York, McGraw-Hill, Inc., 1991.
- REILLY, William J., *Method for the Study of Retail Relationships*, Reserch Monograph nº 4 – Austin: University of Texas Press, 1929. University of Texas Bulletin nº2944)
- ROSEMBLOOM, R, "Trade Area Mix and Retailing Mix", *Journal or Marketing*, September 1976, pg. 58-65.

NOTAS:

¹ BENNETT, Peter D.. (org.). *Dictionary of Marketing Terms*, 2nd ed.. Chicago, American Marketing Association, 1995, p. 287. In: BERMAN, Barry; EVANS, Joel R.. *Retail Management: A Strategic Approach*. Upper Saddle River, Prentice Hall, 1998, p. 276.

² CRAIG, C. Samuel; GHOSH, Avijit; McLAFFERTY, Sara. "Models of the Retail Location Process: A Review". *Journal of Retailing*, Vol. 60, no.1, Spring 1984, p. 15.

Fórmula:

$$S_p = \sqrt{\frac{pxq}{n}}$$

Onde:

s_p = desvio padrão amostral

p = proporção da população dentro de certo raio (2 km)

q = 1-q

n = tamanho da amostra

⁴ APPLEBAUM, William, “Methods for Determining Store Trade Areas, Market Penetration and Potential Sales”, *Journal of Marketing Research*, May, 1966, p. 132

⁵ Modelos gravitacionais de Reilly, Converse.

⁶ De acordo com Bell et al, o modelo de Huff mediria o custo fixo para se fazer as compras, ou seja, mediria o custo do deslocamento ao local de compra, qualquer que fosse a compra. Eles chamam os custos variáveis aqueles relacionados às quantidades de produtos compradas em cada compra e às expectativas de seus preços respectivos. BELL, David R; HO, Teck-Hua; TANG, Christopher S. “Determining where to shop: Fixed and variable costs of shopping”, *Journal of Marketing Research*, Aug 1998 p: 355.

⁷ Escolhemos o hectare por ser uma medida mais intuitiva e de mais fácil entendimento. Um hectare tem 10.000 m² (100 metros de comprimento e 100 metros de largura), e equivale, aproximadamente, a um quarteirão.

⁸ PINDYCK, Robert S.; RUBINFELD, Daniel L.. *Econometric Models & Economic Forecasts*. 3rd. Ed.. New York, McGraw-Hill, Inc., 1991, p. 73.