



Análise do impacto visual da montra no Comportamento do Consumidor utilizando Ferramentas de Data Mining

Cristina Isabel Moreira da Silva

Dissertação de Mestrado em Análise de Dados e Sistemas de Apoio à Decisão

Faculdade de Economia

Universidade do Porto

2009

Orientada por
Prof. Doutor Alípio Jorge
Prof. Doutor Pedro Quelhas Brito

NOTA BIOGRÁFICA

Cristina Isabel Moreira da Silva nasceu em Viana do Castelo em 1980. Licenciou-se em Economia na Faculdade de Economia do Porto em 2003.

Em 2004, obteve a competência de Técnica Oficial de Contas, e ainda neste ano iniciou a sua actividade profissional num grupo têxtil, na área de controlo de gestão.

Em 2006 obteve a licenciatura em Gestão na mesma faculdade. No ano lectivo 2007/2008 completou a parte escolar do Mestrado de Análise de Dados e Sistemas de Apoio à Decisão, na Faculdade de Economia do Porto.

Actualmente encontra-se a colaborar com uma empresa que tem por actividade o comércio por grosso e a retalho, enquanto responsável administrativa e financeira da mesma.

AGRADECIMENTOS

Aos orientadores desta dissertação, o Professor Doutor Alípio Jorge e Professor Doutor Pedro Quelhas Brito, pelos conselhos e sugestões fornecidos ao longo deste percurso.

Aos responsáveis da Salsa e da Trotleman pela colaboração, interesse e disponibilidade demonstrados, bem como ao António Manuel Sampaio Ribeiro por disponibilizar a base de dados que construiu com base nos dados recolhidos.

Aos meus familiares, namorado e amigos, pelo apoio e força que me deram.

RESUMO

A Análise de Dados e a Extracção de Conhecimento de Dados são ferramentas que têm ganho cada vez maior relevância, dado que a quantidade de dados que temos à disposição é cada vez maior e a importância que estes dados assumem na tomada de decisão é crucial.

O presente trabalho tem assim por objectivo aplicar a análise e extracção de dados à análise do poder de atracção da montra nos potenciais consumidores. Assim propomo-nos a usar a ferramenta das Árvores de Decisão, para analisar em que medida o comportamento dos consumidores é ou não influenciado pela montra, isto é, em que medida o consumidor entra na loja e compra em resultado do poder de atracção da montra. Como os dados em estudo foram recolhidos na época de saldos e na época da nova colecção propomo-nos também a analisar em que medida o poder de atracção da montra e a sua influência no comportamento do consumidor é diferente consoante se está ou não na época de saldos.

Pretende-se com este trabalho contribuir para um melhor conhecimento deste instrumento de comunicação que todos os retalhistas têm à sua disposição, e assim contribuir para que retirem da montra a maior rentabilidade possível.

ABSTRACT

Data Mining and Knowledge Discovery from Data are tools that have attracted a great deal of attention in recent years, since every day the amount of data stored increase and tall his data is crucial to management decisions.

The goal of the present dissertation is to apply the data mining and the Knowledge Discovery from Data to the analyses of the power that a store window may have to pick up potential buyers. So this work will use the classification to find a model that describes and distinguishes data classes. This model will be represented in the form of decision trees. With this method we will be able to analyze how the behavior of the potential buyers is a consequence of the store window: we want to know if a person goes into the store and buy products as a consequence of the influence of the store window.

The data that will be analyzed was collected in the sales season and when the store already presented products of the new season (Fall-Winter). So we also pretend to analyze if the store window as different influence in the potential buyer in this two different moments of time, were different collections are presented in the store.

With this Job we will contribute for a better knowledge of the store window as an important instrument of communication, available for all managers of the retail industry, and so help them to obtain the best profitability possible with this instrument.

ÍNDICE

Nota Biográfica.....	i
Agradecimentos	ii
Resumo	iii
Abstract.....	iv
Índice	v
1. Introdução	1
1.1. Motivações e Contribuições.....	1
1.2. Definição do Problema.....	2
1.3. Estrutura da Dissertação.....	3
2. Estado da Arte: O Comportamento dos Consumidores - Perspectiva de Gestão.....	6
3. Estado da arte: Análise de Dados e Data Mining.....	12
3.1. As Árvores de Decisão.....	14
3.2. Porquê Árvores de Decisão.....	17
3.3. Aplicação das Árvores de Decisão.....	18
4. Análise Exploratória de Dados	22
4.1. Apresentação das Variáveis	22
4.2. Recolha de dados.....	27
4.3. Análise Preliminar.....	28
4.4. Análise Descritiva das Variáveis	34

4.4.1.	Análise do Comportamento das Variáveis.....	34
4.4.1.1.	Equilíbrio da Amostra	35
4.4.1.2.	Deslocação no Corredor.....	36
4.4.1.3.	Comportamento face à montra	36
4.4.1.4.	Caracterização dos Indivíduos /Grupos	39
4.4.2.	Afluência dos Indivíduos	43
4.4.3.	Decisão de Compra.....	45
5.	Data Mining	47
5.1.	Quem Entra	48
5.1.1.	Modelo de classificação obtido para a variável “Entra ou Não na Loja” no dia 9 de Setembro	48
5.1.2.	Modelo de classificação obtido para a variável “Entra ou Não na Loja” no dia 16 de Setembro	52
5.2.	Quem Compra	54
5.2.1.	Modelo de classificação obtido para a variável “Compra ou Não” no dia 9 de Setembro	55
5.2.2.	Modelo de classificação obtido para a variável “Compra ou Não” no dia 9 de Setembro	58
6.	Discussão dos resultados e aplicação no negócio	63
7.	Conclusão.....	71
7.1.	Resultados	71
7.1.1.	Qual a influência da montra na decisão de entrar ou não na loja?.....	72
7.1.2.	Quais as variáveis mais importantes na decisão de compra?.....	75
7.1.3.	A decisão de entrar e de comprar é ou não influenciada por variáveis diferentes consoante estamos ou não na época de saldos?	78
7.2.	Principais Contribuições	79
7.3.	Limitações e Trabalhos Futuros	81
8.	Bibliografia	83
9.	Anexos	86

ÍNDICE DE TABELAS

Tabela 1: Diagrama de uma matriz de confusão.....	19
Tabela 2: Apresentação das Variáveis	23
Tabela 3: Estatísticas Descritivas das Variáveis Qualitativas	29
Tabela 4: Estatísticas Descritivas das Variáveis Quantitativas	30
Tabela 5: Mediana das Variáveis em Estudo.....	31
Tabela 6: Desempenho da árvore de decisão em relação à variável “Entra ou não na Loja” no dia 9 de Setembro	49
Tabela 7: Desempenho da árvore de decisão em relação à variável “Entra ou não na Loja” no dia 16 de Setembro	53
Tabela 8: Desempenho da árvore de decisão em relação à variável “Compra ou não” no dia 9 de Setembro	56
Tabela 9: Desempenho da árvore de decisão em relação à variável “Compra ou não” no dia 16 de Setembro	59
Tabela 10: Variáveis de cada nível de cada uma das Árvores de Decisão “Entra ou Não na Loja” do dia 9 e 16 de Setembro.....	74
Tabela 11: Variáveis de cada nível de cada uma das Árvores de Decisão “Compra ou Não” do dia 9 e 16 de Setembro	77

ÍNDICE DE GRÁFICOS

Gráfico n.º 1: Box Plot do Tempo de Observação da Montra	32
Gráfico n.º 2: Box Plot do Tempo na Loja	32
Gráfico n.º 3: Box Plot do Número de Sacos de Compras	33
Gráfico n.º 4 Box Plot do Número de Indivíduos que compõe cada Observação	33
Gráfico n.º 5: Distribuição dos dados pela variável “Entra ou Não Entra na Loja”	35
Gráfico n.º 6: Distribuição dos dados pela variável “Número de Sacos de Compras” ...	36
Gráfico n.º 7: Distribuição dos dados pelas variáveis “Passa ou Não Passa em frente da montra” e “Deslocação no interior/exterior do corredor”	37
Gráfico n.º 8: Distribuição dos dados pelas variáveis “Ataque à montra” e “Deslocação no Corredor”	38
Gráfico n.º 9: Distribuição das observações pelas variáveis “Observação da montra” e “Comportamento em Relação à montra”	39
Gráfico n.º 10: Distribuição dos dados pelas variáveis “Número de indivíduos” e “Número de Sacos de Compras”	40
Gráfico n.º 11: Proporção de indivíduos que entra ou não na loja, por idade, nos dois dias observados	41
Gráfico n.º 12: Número de compras por grupo de idade, nos dois dias de observação ..	41
Gráfico n.º 13: Proporção de Observações face à variável Entra ou não na loja, por sexo, por dia de observação	42
Gráfico n.º 14: Compra por sexo por dia de observação	42
Gráfico n.º 15: Composição dos grupos constituídos pelos 2 sexos que realizam compras	43
Gráfico n.º 16: Número de observações por período do dia	44

Gráfico n.º 17: Número de Observações em que se entra na loja por período do dia por dia de observação 44

Gráfico n.º 18: Número de observações em que se compra por período do dia por dia de observação 44

ÍNDICE DE FIGURAS

Figura n.º 1: Representação Gráfica de uma Árvore de Decisão.....	15
Figura n.º 2: Árvore de Classificação “Entra ou não na loja” no dia 9 de Setembro	49
Figura n.º 3: Árvore de Decisão “Entra ou não na loja” no dia 16 de Setembro.....	52
Figura n.º 4: Árvore de Decisão “Compra ou não” no dia 9 de Setembro	56
Figura n.º 5: Árvore de Decisão “Compra ou não” no dia 16 de Setembro	59

ÍNDICE DE ANEXOS

Anexo n.º 1: Comportamento face à montra.....	86
Anexo n.º 2: Output gerado pelo Weka para o modelo das Árvores de Decisão em relação ao atributo “Entra ou não na Loja” no dia 9 de Setembro.....	87
Anexo n.º 3: Output gerado pelo Weka para o modelo das Árvores de Decisão em relação ao atributo “Entra ou não na Loja” no dia 16 de Setembro.....	89
Anexo n.º 4: Output gerado pelo Weka para o modelo das Árvores de Decisão em relação ao atributo “Compra ou não” no dia 9 de Setembro	91
Anexo n.º 5: Output gerado pelo Weka para o modelo das Árvores de Decisão em relação ao atributo “Compra ou não” no dia 16 de Setembro	93

1. INTRODUÇÃO

1.1. MOTIVAÇÕES E CONTRIBUIÇÕES

O presente trabalho tem por objectivo analisar, a partir de um conjunto de dados, em que medida o comportamento de um consumidor face a uma montra nos permite aferir se esse consumidor entra ou não na loja e se compra ou não, e em que medida o comportamento do consumidor é diferente consoante estamos ou não na época de saldos.

O sucesso de um retalhista pode ser medido pela sua capacidade de conseguir, a longo prazo, obter um desempenho financeiro superior face aos concorrentes através da adaptação dos seus recursos às oportunidades de mercado. Assim atrair e manter os consumidores constituem os principais objectivos do retalhista, pois só assim é possível tornar a actividade rentável [1]. Diversos instrumentos estão disponíveis para atingir este objectivo, no entanto, o que se pretende aqui analisar é em que medida a montra pode ou não ser um importante instrumento para captar a atenção dos consumidores, incentivando-os a entrar na loja e a comprar.

A compra é o resultado de um processo cognitivo de cada consumidor, relacionado com os seus comportamentos e a sua personalidade, bem como da sua interacção com outros indivíduos, e da envolvente física e social do local onde se encontra a loja [2]. Nesta envolvente física e social onde a loja se situa, o consumidor tem consciência da loja através da montra, pois este é o seu primeiro contacto com ela. Ele poderá conhecer a marca ou os artigos que estão à venda, mas é a montra que lhe dá uma primeira impressão da loja em si. De facto, temos assistido a um progressivo aumento da importância da actividade retalhista face aos distribuidores grossistas, uma vez que possuem maior visibilidade, pelo facto de interagirem directamente com o consumidor, nomeadamente através da montra [3]. A importância da montra é um facto e uma realidade presente no mercado, daí que o valor de uma loja esteja directamente relacionado com a dimensão da montra e com a própria localização da loja [2] [3].

O conjunto de dados em estudo foi recolhido no Norte Shopping, mais especificamente na loja Salsa: uma câmara de vídeo exterior permitiu captar os consumidores que se deslocavam no exterior da loja, e que paravam ou não para ver a montra; uma câmara de vídeo interior permitiu analisar se os consumidores realizavam ou não compras. As imagens foram recolhidas no dia 9 e 16 de Setembro de 2006, na época de saldos e na época de nova colecção, respectivamente. Posteriormente, o comportamento registado dos consumidores foi manualmente transcrito para um conjunto de dados no âmbito do trabalho de dissertação de António Ribeiro [21]. Este conjunto de dados foi o objecto directo da nossa análise.

1.2. DEFINIÇÃO DO PROBLEMA

O presente trabalho tem por objectivo analisar em que medida o comportamento de um consumidor face a uma montra nos permite aferir se esse consumidor entra ou não na loja e se compra ou não, consoante estamos na época de saldos ou na época da nova colecção.

Este problema pode ser definido em três pontos distintos:

1. Avaliar a influência da montra na decisão de entrar ou não na loja. Pretende-se analisar em que medida o indivíduo entra ou não na loja em função do comportamento no corredor, da deslocação no corredor, do facto de passar ou não em frente da montra, de olhar ou não para a montra, do ataque à montra, do facto de parar ou não em frente da montra, do movimento em frente à montra, do comportamento face à montra, e da observação da montra;
2. Relacionar a decisão de compra com o comportamento do indivíduo face à montra e assim averiguar quais as variáveis mais importantes na decisão de compra;
3. Analisar em que medida a decisão de entrar e de comprar é ou não influenciada por factores distintos, consoante estamos ou não na época de saldos.

Pretende-se obter resultados que forneçam recomendações de como otimizar a montra enquanto veículo de comunicação e atracção dos consumidores. De facto, esta é o primeiro contacto do indivíduo com uma loja e como tal é a que lhe permite ficar com uma primeira impressão da mesma, pelo que é um instrumento muito importante [2].

Para a análise deste problema, que tem por objectivo aferir da informatividade do comportamento do consumidor, será usado o método de indução de Árvores de Decisão. A adopção deste método resulta do facto de permitir explicar a decisão de entrar ou não na loja e de comprar ou não, consoante é ou não época de saldos. Assim serão construídas Árvores de Decisão, cujo conceito é apresentado no Capítulo 3, usando como valores de classe as variáveis “Compra” ou “Não Compra”, para responder à questão colocada no ponto 1; e “Entra” ou “Não Entra” para responder à questão colocada no ponto 2. As árvores serão construídas usando dados que descrevem o comportamento dos consumidores referentes à época de saldos e à época da nova colecção, para se averiguar se as conclusões obtidas são distintas consoante o dia de observação e, desta forma, tentar apurar o impacto da época de saldos.

Para a melhor compreensão de como está estruturada a presente dissertação é de seguida apresentada a sua estrutura.

1.3. ESTRUTURA DA DISSERTAÇÃO

Quanto à estrutura desta dissertação, numa primeira fase pretendeu-se mostrar qual a motivação subjacente ao presente trabalho, bem como as principais contribuições e definir claramente o problema que este trabalho se propõe a estudar, que consiste em analisar o impacto da montra no potencial consumidor, e avaliar em que medida a decisão do consumidor de entrar e comprar é influenciada pela montra.

Posteriormente será feita uma breve análise do comportamento dos consumidores, numa perspectiva de gestão, pois pretende-se explorar os conceitos e definições subjacentes aos dados que estão a ser objecto de análise, de modo a se poder tirar conclusões face

aos comportamentos registados e se saber como intervir nesses comportamentos para se gerar a compra. Assim, pretende-se responder a questões como qual a importância da montra, da marca, da loja e das promoções. Pretende-se com este capítulo transmitir uma noção dos conceitos subjacentes à análise que se está a realizar.

De seguida é apresentado a Análise de Dados e a Extração de Conhecimentos de Dados (Data Mining), por vezes aqui abreviado para ECD, como uma importante ferramenta para a análise do problema proposto. De entre os métodos disponíveis é apresentado o método escolhido, um método de indução das Árvores de Decisão, e as razões que motivaram a sua escolha. Este método será aplicado através do programa Weka, que também será brevemente exposto.

No sentido de ganhar sensibilidade ao problema e de averiguar a coerência dos dados é realizada uma análise exploratória dos mesmos. Numa primeira fase, foram apresentadas e descritas as variáveis em estudo, bem como o método de recolha dos dados. Posteriormente, efectuou-se uma análise preliminar dos dados, seguida da análise de algumas estatísticas descritivas e de localização. Por fim, é apresentada uma análise descritiva das variáveis, com o objectivo de compreender o comportamento das mesmas, uma vez que para a aplicação do método e compreensão dos resultados obtidos será importante ter uma visão geral das variáveis e do seu comportamento, a qual é assim apresentada neste capítulo 4.

No capítulo seguinte é aplicado o método de ECD escolhido aos dados em causa, com o objectivo de analisar quem entra e quem compra antes e após a época de saldos. São assim expostos os modelos e as conclusões obtidas, tendo-se constatado que a maioria dos indivíduos é classificada como entra na loja sem observar a montra e que a maioria dos indivíduos é classificada como compradora caso permaneça na loja por um certo período de tempo, consoante o sexo e o número de indivíduos que fazem parte da observação do dia 9 e 16 de Setembro, respectivamente.

Os resultados serão objecto de discussão e tentar-se-á aplicar as conclusões obtidas ao negócio em causa, no sentido de contribuir para o uso da montra como importante instrumento de rentabilização da actividade do retalhista, ao captar clientes e incentivar a compra.

Por fim, todo o trabalho e as conclusões obtidas serão apresentadas e explicadas a título de conclusão, bem como apresentadas as contribuições deste trabalho para a área em estudo e os futuros trabalhos a realizar. Espera-se assim contribuir para o estudo da montra, enquanto instrumento disponível a todos os comerciantes, de modo a se conseguir obter deste instrumento de comunicação e captação dos consumidores, o máximo retorno.

2. ESTADO DA ARTE: O COMPORTAMENTO DOS CONSUMIDORES - PERSPECTIVA DE GESTÃO

O objectivo de qualquer empresa é conseguir vender os seus produtos e serviços e, desta forma, obter liquidez para fazer face aos seus compromissos, tendo como meta a obtenção do lucro [1]. Para tal a imagem da empresa e dos produtos e/ou serviços que oferece junto dos consumidores é essencial [2]. Esta imagem é veiculada através de diferentes instrumentos, dos quais se destaca a marca.

De facto, a marca é a “cara” de uma empresa perante o mundo, representando os produtos e serviços que comercializa de uma forma distinta das outras empresas [4]. A força de uma marca está relacionada com a qualidade percebida dos produtos que oferece, com o seu conhecimento pelos consumidores, com a posição de mercado que detém e com a sua identidade; e traduz-se na lealdade, na disposição dos consumidores para pagarem um maior preço e para aceitarem outras extensões da marca [4] [5]. Daqui se retira que o poder da marca sobre o consumidor é um importante factor a ter em conta na análise da decisão de compra, e que se tem de ter presente.

Por outro lado, a montra é um veículo de transmissão dessa informação para o consumidor, e é o primeiro contacto que um transeunte tem com a loja, e consequentemente com a empresa [2] [3]. A montra é o cartão-de-visita de uma loja [6]. A decisão de entrar ou não na loja, essencial para que possa haver compra, pode ser motivada por diferentes factores, entre os quais a observação da montra. É assim importante avaliar em que medida ela influencia a decisão do consumidor de entrar na loja e comprar.

A montra objecto de estudo pertence a uma loja localizada num centro comercial, pelo que o impacto da montra e a captação dos potenciais clientes está inserida numa problemática mais global que tem a ver com a deslocação dos potenciais consumidores nos centros comerciais, a qual tem sido analisada por diversos autores [2].

Os modelos espaciais de comportamento, tais como Huff ou Cadeia Markov são eficientes a estimar o potencial de atracção de um novo centro comercial, ou a prever o

número de visitantes, ou que um determinado local será provavelmente escolhido por alguma entidade, devido ao número de potenciais cliente que pode abranger. No entanto, não conseguem lidar com o movimento individual e pedestre de um indivíduo, numa escala menor, como os centros comerciais [7].

Neste sentido, foram construídos modelos multi-agente para investigar os movimentos pedestres, nomeadamente, o comportamento dos consumidores em relação aos centros comerciais [7]. Para tal, observou-se a rota seguida por diferentes consumidores em diferentes centros comerciais, e foram realizados inquéritos acerca dos seus perfis quanto às suas preferências e os seus objectivos, para se tentar averiguar quais os destinos mais prováveis.

Esta análise permitiu classificar os consumidores em quatro categorias: aqueles que fixam o destino no início da rota pelo centro comercial e seguem o mesmo percurso quase todos os dias; aqueles que fixam destinos e rotas que variam nas diferentes idas aos centros comerciais; aqueles que não definem qualquer rota mas procuram determinados produtos; aqueles que não fixam nenhuma rota mas gostam de window shopping sem qualquer objectivo.

Neste estudo, a comparação dos resultados da simulação com os reais demonstrou que o comportamento espacial dos consumidores é constituído por vários segmentos, em que o consumidor, em cada segmento que percorre, rege-se pelo princípio da utilidade máxima, no entanto, a rota como um todo não pode ser explicada, segundo esse princípio. Porém, a importância de classificar um potencial consumidor em relação ao seu objectivo, numa deslocação ao centro comercial é indiscutível, uma vez que o permite conhecer e assim definir estratégias no sentido de o tornar um consumidor efectivo.

Numa outra abordagem a este problema procurou-se investigar a representação mental dos consumidores acerca das decisões de rota nos centros comerciais, e de como podem variar dependendo do contexto [8]. Esta análise teve em consideração o impacto do contexto do centro comercial, nomeadamente de factores como a localização, o tempo e a acessibilidade de transportes. Procurou analisar se diferentes contextos, em termos de ambientes de compra, levam à activação de diferentes atributos devido à percepção pelo consumidor de diferentes benefícios; e se variações de contexto têm consequência na

variação da utilidade dos consumidores devido a alterações na percepção dos atributos e benefícios.

Concluiu-se que, se forem conhecidos quais os benefícios que o consumidor percebe, se conseguem obter melhorias na previsão dos atributos relevantes, e que em diferentes contextos, são accionados diferentes benefícios, percebidos por diferentes atributos. Por outro lado, também se concluiu que diferentes benefícios e atributos, de facto levam a diferentes avaliações das alternativas.

Apesar da importância que tem a questão de qual o percurso que um consumidor realiza num centro comercial, pois dita o seu sucesso ou fracasso, bem como das lojas que nele se situam, muito pouco tem sido feito nesta área. A principal dificuldade reside na compreensão dos estudos realizados nesta matéria. Uma variedade de técnicas está disponível mas cada uma é ineficiente em alguns aspectos. Os questionários muitas vezes obtêm respostas não verdadeiras, a observação, a contagem do número de passagens é uma alternativa que não consegue captar o movimento individual mas apenas de grupos. Já a gravação permite observar o que as pessoas fazem mas não a razão de o fazerem [9].

Foi realizado um estudo, durante uma semana num centro comercial, em que se observou 250 grupos de indivíduos, desde a entrada no centro comercial até à sua partida, tendo-se verificado o seguinte [9]:

- Destes 250 grupos, oito não estavam no centro comercial para realizar compras;
- A percentagem de indivíduos que entra e compra é mais elevada nos serviços, à excepção das agências de viagens;
- Dos grupos observados, 11% circularam através de todo o centro comercial; 1/3 restringiu os movimentos a menos de metade do espaço disponível no centro comercial, 2/3 dos grupos entraram em lojas, mais de 90% das compras foram realizadas em três ou menos secções do centro comercial. A zona do centro comercial mais visitada foi a intercepção entre os corredores centrais este-oeste e norte-sul;

- O típico grupo de 2 pessoas entra em média em 5 lojas e realiza compras em 3 delas, permanecendo no centro comercial cerca de 50 minutos;
- O estudo demonstrou ainda que mesmo com 180 000 m² a grande maioria dos indivíduos restringe-se a uma pequena parte do centro comercial.

Assim o que motiva o consumidor a seguir um dado percurso? A resposta a esta questão está dependente da motivação de um consumidor para se deslocar a um centro comercial, do seu objecto e em última instância do facto de gostar ou não de o fazer. Os consumidores que gostam de visitar os centros comerciais realizam maior número de compras não planeadas e passam mais tempo às compras e continuam a comprar depois de realizarem uma compra. Estes compradores gostam de visitar novas lojas. Assim é óbvio que os retalhistas pretendem atrair estes consumidores às suas lojas. Alguns investigadores argumentam que este tipo de consumidores prefere window shopping mais do que qualquer outro [10].

Mas assim sendo como captar estes consumidores e leva-los a comprar? Um importante meio é sem dúvida a montra. Se prever qual a rota que o consumidor irá realizar num centro comercial é algo extremamente difícil, assim como qual o centro comercial que irá escolher, uma importante estratégia é captar a atenção de todos aqueles que por algum motivo decidem passar em frente da loja.

Assim passamos a estar no domínio da relação da loja e da marca com o consumidor e é aqui que pretendemos dar alguns passos. O conhecimento de quais os aspectos da montra que captam mais a atenção do consumidor traduzir-se-ia numa importante vantagem competitiva, porque permitiria aumentar as vendas e, consequentemente, a produtividade da loja. No limite, saber como promover um artigo e como vender pode ser muito importante a nível inclusive da gestão de stocks [11]. Assim a montra é uma peça de gestão fundamental, que merece a atenção do retalhista.

Um outro aspecto que é abordado neste trabalho tem a ver com os saldos, uma vez que a amostra é recolhida em dois períodos: na época de saldos e na nova colecção. Interessa assim analisar este instrumento de promoção de vendas. As promoções de vendas são acontecimentos de marketing focalizados na acção cujo propósito é produzir um impacto directo no comportamento dos clientes das empresas [12]. O uso das

promoções para influenciar o comportamento do consumidor tem aumentado progressivamente em resultado do poder cada vez maior dos retalhistas, da diminuição da lealdade à marca, do aumento da sensibilidade dos consumidores às promoções, do número cada vez maior de novos produtos e da fragmentação mercado [13].

Inquéritos realizados aos consumidores procuraram determinar se e quanto as promoções aumentavam a escolha por uma dada marca. Também se analisou se as promoções têm um impacto que se prolonga para além período em que decorrem. As promoções podem aumentar a preferência por uma dada marca aos consumidores que já a usam, através do efeito que provocam ao lembrar os consumidores de a comprarem. Por outro lado, podem induzir os consumidores que não a usam, a experimentar [14]. Assim o uso das promoções poderá ser considerado excessivo quando existe uma tal dependência deste instrumento que se está disposto a sacrificar a posição e a imagem da marca a longo prazo pelo aumento das vendas no curto prazo [13].

Neste estudo foram identificados dois indicadores para avaliar se as promoções alteram a preferência pela marca após terminarem: percepção da marca e probabilidade de escolha. O objectivo foi analisar o impacto de oferecer versus não oferecer promoções; e de oferecer promoções com maior ou menor intervalo de tempo. Foram detectadas três variáveis que influenciam os consumidores na altura das promoções e que consequentemente podem afectar o seu comportamento na pós-promoção [14]:

- Características da promoção: se é um cupão, a oferta de um prémio, valor;
- Características do produto: frequência de compra, se o produto é uma experiência boa, o nível de preço da categoria dos produtos em causa; o nível de preços da marca dentro dessa categoria, a popularidade da marca;
- Características dos consumidores.

Concluiu-se que em média as promoções não afectam a preferência pela marca depois de terminar o período promocional, no entanto, tal pode acontecer em determinadas condições. Por outro lado concluiu-se que as características dos consumidores não parecem afectar o comportamento após a época de promoções, enquanto as características dos produtos e da promoção afecta a preferência pela marca.

Esta é assim uma área em que não existe consenso quanto aos eventuais efeitos que esta técnica de vendas pode ter no longo prazo, sendo que no curto prazo é inegável a sua capacidade de aumentar o volume de vendas face ao efeito atractivo que tem nos consumidores [15].

Este trabalho propõe-se a analisar em que medida se pode ou não distinguir padrões de comportamento diferentes consoante a observação da montra ocorre ou não num dia de saldos e, em que medida essa alteração de comportamento se traduz ou não num maior número de entradas na loja e num maior volume de compras pelos consumidores.

Uma vez realizada uma síntese da problemática em torno da marca e da sua visibilidade, da montra, da loja e da sua localização e do efeito das promoções, vamos de seguida debruçar-nos sobre a problemática em torno da Análise de Dados, a qual se coloca face aos dados que estão a ser objecto de estudo e através dos quais se pretende responder às questões colocadas.

3. ESTADO DA ARTE: ANÁLISE DE DADOS E DATA MINING

Actualmente, dada a facilidade que existe em armazenar e guardar dados, há uma enorme quantidade de dados disponível acerca dos mais variados assuntos. [16] Assim, estão ao dispor de todos as mais variadas bases de dados, que aumentam de dia para dia, mas qual o nosso conhecimento acerca destes dados?

A verdade é que, à medida que a quantidade de informação disponível aumenta, a proporção de conhecimento que as pessoas detêm diminui, porque por detrás de todos estes dados estão informações relevantes que raramente estão explícitas. O desafio que se coloca é então encontrar um sentido no conjunto de dados analisados, e a partir da descoberta de padrões relevantes, obter conhecimentos desses dados e usar esse conhecimento para prever o que irá acontecer em novas situações [16].

A Extração de Conhecimento de Dados pretende fazer face a este objectivo, isto é, analisar os dados existentes, utilizando um conjunto de instrumentos, e assim resolver problemas utilizando os padrões descobertos nesses dados para realizar previsões e obter informação que pode ser relevante na tomada de decisão. Segundo Jiawei e Kamber podemos definir Data Mining como a extração de padrões de conhecimento implícitos em grandes bases de dados, pelo que deveria ter sido denominada de uma forma mais própria de extração de conhecimento de dados. Mining caracteriza o processo de procura de informação preciosa de entre um grande volume de informação [17].

Para alguns Data Mining é sinónimo de extração de conhecimento de dados, enquanto para outros é uma fase desse processo de descoberta de conhecimento, o qual é composto por um conjunto de fases interactivas: limpeza dos dados, combinação de dados que possam ser juntos, selecção dos dados, transformação dos dados, Data Mining, avaliação dos padrões encontrados e finalmente o conhecimento [17].

Tipicamente as bases de dados usadas para aplicar técnicas de Data Mining são compostas por um conjunto de exemplos (observações). Cada exemplo é descrito por

um conjunto de atributos (variáveis). Cada atributo tem um conjunto de valores, podendo estes ser categóricos ou contínuos [16].

O Data Mining, pode ser classificado como descritivo, quando o objectivo é o de encontrar padrões que caracterizem os dados, ou como preditivo, quando se pretende utilizar a informação escondida em dados para realizar previsões [17].

Dentro do Data Mining preditivo, os tipos de modelos mais importantes são os de classificação e os de previsão numérica. Nos modelos de classificação o objectivo é descrever e distinguir as classes para um dado conjunto de dados, com o objectivo de prever a classe de um novo exemplo. Este processo tem duas fases: primeiro o modelo é construído a partir de um conjunto de dados onde um determinado atributo é o atributo classe (atributo objectivo). Esta é a fase de treino ou de aprendizagem. Posteriormente o modelo é usado para classificar um novo exemplo para o qual a classe é desconhecida. A aprendizagem do modelo realiza-se para um conjunto de dados, enquanto a aplicação do modelo é realizada noutros dados. Quando se pretende estimar o desempenho do modelo, utiliza-se um conjunto de dados de teste. Frequentemente o desempenho do modelo é expresso em termos da percentagem de casos de teste que são correctamente classificados. Existem muitos métodos para obter modelos de classificação a partir de dados, tais como métodos de indução de Árvores de Decisão, métodos Bayesianos e o Discriminante Linear [17].

Os modelos de previsão numérica são usados para prever valores desconhecidos de um determinado atributo numérico. Também é um processo com duas fases semelhantes às anteriormente descritas, no entanto neste caso estamos perante uma variável objectivo contínua e não categórica como acontece nos modelos de classificação. O método de previsão numérica mais utilizado é a regressão [17].

No nosso caso iremos usar um método de classificação porque as variáveis objectivo são categóricas. Em particular trata-se de problemas de duas classes. Vamos então analisar com mais pormenor alguns dos métodos que temos à nossa disposição.

Os classificadores Bayesianos prevêem a classe de um caso de acordo com a probabilidade de esse caso pertencer a uma dada classe. Este método tem por base o Teorema de Bayes segundo o qual um caso pertencerá a uma dada classe de acordo com a probabilidade condicionada de esse caso pertencer a essa classe, dados os valores que

são conhecidos para os restantes atributos. O Discriminante Linear utiliza a função densidade de probabilidade no cálculo das probabilidades de modo a que cada caso seja classificado na classe que maximiza essa probabilidade. As Árvores de Decisão formulam regras de classificação para os dados de acordo com o ganho de informação com que cada atributo contribuiu [16].

As Árvores de Decisão permitem também caracterizar os dados em causa, o que no caso em estudo é muito importante pois permitirá descrever padrões de comportamento dos consumidores, e simultaneamente permite usar essas conclusões para prever novos casos. Este foi, portanto, o método escolhido, pelo que é de seguida apresentado com maior detalhe.

3.1. AS ÁRVORES DE DECISÃO

As Árvores de Decisão têm-se revelado um método eficaz para abordar problemas de classificação, pois permitem trabalhar com grandes quantidades de dados e a representação é intuitiva e fácil de interpretar [17]. Este foi o método escolhido para a análise dos dados nesta dissertação.

Uma árvore de decisão, tem um aspecto gráfico de uma árvore, iniciando-se na raiz e posteriormente alargando os seus ramos. É composta por nós, ramos e folhas: cada nó é composto por um dado atributo, cada folha é composta por uma classe, cada ramo dá origem a um nó ou uma folha, e assume um dado valor para o atributo presente no nó que lhe deu origem. Assim na representação das Árvores de Decisão (Figura n.º 1) [17]:

- Cada nó de decisão contém um teste num atributo.
- Cada ramo descendente corresponde a um possível valor deste atributo.
- Cada folha está associada a uma classe.
- Cada percurso na árvore (da raiz à folha) corresponde a uma regra de classificação.

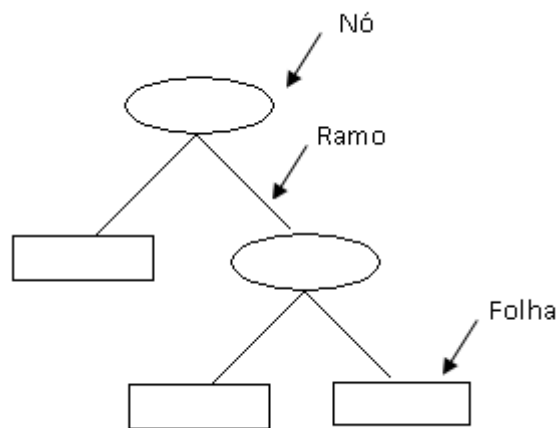


Figura n.º 1: Representação Gráfica de uma Árvore de Decisão

Uma questão que se coloca é então saber quais os atributos que são seleccionados para fazer parte da árvore de decisão. Para tal são usadas medidas de selecção dos atributos que pretendem escolher o melhor critério de separação dos dados. De facto estas medidas fornecem uma hierarquia dos atributos em função da classificação que fazem dos exemplos. O atributo que tiver o melhor desempenho é o escolhido para separar os dados [17].

Uma outra questão relevante que se coloca ao modelo das Árvores de Decisão é o “overfitting” dos dados, isto é, o modelo obtido poderá descrever de tal forma exacta os dados que lhe dão origem que se pode revelar ineficaz para prever um novo caso. Assim, para evitar este problema são usados métodos de poda da árvore de decisão, que utilizam medidas estatísticas para remover os ramos das árvores com menor nível de confiança. As Árvores de Decisão podadas tendem a ser menores, menos complexas e logo mais fáceis de compreender, e são normalmente mais rápidas e melhor a classificar um novo exemplo [17].

O desempenho dos modelos obtidos através de um método de indução de árvore de decisão pode ser medido através da taxa de erro, isto é, da percentagem de exemplos que os modelos classificam incorrectamente. Para aferir esta taxa de erro podem ser usados os dados utilizados para construir o modelo (dados de treino), ou então um conjunto de dados separado, usados apenas para testar o modelo (dados de teste). Esta segunda opção permite estimar a taxa de erro dos modelos de forma mais fiável [17].

As Árvores de Decisão podem ser obtidas através de diversas ferramentas, entre as quais podemos referir o programa R e o Weka (Waikato Environment for Knowledge Analysis). Nesta dissertação será utilizada o Weka, que permite a aplicação de métodos de aprendizagem a um conjunto de dados de um modo amigável, disponibilizando um conjunto de funcionalidades adicionais.

Assim, no caso deste programa, um dos métodos de indução de Árvores de Decisão disponíveis é o J48, que é uma versão do algoritmo C4.5 de Ross Quinlan [18]. Este método usa para a selecção dos atributos que compõem a árvore de decisão, a medida do ganho de informação. Ou seja, para cada nó é seleccionado o atributo que permite obter um maior ganho de informação com a partição dos exemplos pelos valores desse atributo. No cálculo desta medida é usado o conceito de entropia de informação, sendo que esta medida assume o valor zero quando todos os exemplos pertencem à mesma classe, ou seja não existe qualquer ganho de informação em seleccionar um atributo para repartir os dados. É obtida pela seguinte expressão:

$$\text{Entropia}(p_1, p_2, \dots, p_n) = -p_1 \log p_1 - p_2 \log p_2 \dots - p_n \log p_n$$

Em que $p_1, p_2, \dots, p_n$ é dado pela proporção de exemplos que pertencem à classe 1, 2, ..., n de um dado atributo. [17]

Para evitar o problema do overfitting do modelo aos dados, este algoritmo usa o método de poda “postpruning”. Este método o que faz é, após a construção da árvore, remover os ramos (e os nós e ramos subsequentes) e substituir os atributos que lhes deram origem por folhas. Esta folha corresponderá à classe mais frequente na sub-árvore que substitui. O critério para realizar a poda da árvore está relacionado com a taxa de erro do modelo nos dados de treino. Assim começando deste a raiz da árvore, em cada nó é calculada a taxa de erro do modelo nesse nó, e tal é comparado com a taxa de erro do modelo se a árvore fosse podada nesse nó, isto é, se o nó fosse substituído por uma folha. Se desta comparação resultar que o menor custo é obtido com a poda da árvore, então a árvore será podada [17].

Após se expor o funcionamento do modelo das Árvores de Decisão será importante perceber o que motiva a opção por esta alternativa, o que será exposto no ponto seguinte.

3.2. PORQUÊ ÁRVORES DE DECISÃO

A adoção de um método de indução de Árvores de Decisão foi motivada pelos seguintes factores [16] [17]:

- Os modelos obtidos são fáceis de interpretar. Assim, para se explicar uma classificação em particular é apenas necessário olhar para as decisões que conduziram a essa classificação;
- É um método não paramétrico: não assume nenhuma distribuição para os dados e pode construir modelos para qualquer função desde que o número de exemplos de treino seja suficiente;
- Permite lidar com variáveis independentes categóricas e contínuas;
- É robusto a variáveis que não têm qualquer valor preditivo: são ignoradas pelo que não afectam a construção do modelo;
- Revela interacções complexas entre as variáveis;
- Cada ramo de uma árvore contém diferentes combinações de variáveis e a mesma variável pode aparecer em diferentes ramos da árvore. Isto pode revelar como uma variável pode depender de outra e em que contexto específico esta dependência existe;
- É robusto à presença de pontos extremos;

Porém é necessário ter consciência que, apesar de o método das Árvores de Decisão ser um poderoso instrumento de Análise de Dados, não é um instrumento perfeito [16] [17]:

- Pode ser instável e uma pequena variação dos dados pode gerar Árvores de Decisão muito diferentes;

- As árvores geradas podem ser muito grandes e complexas, o que pode tornar o modelo difícil de interpretar, de compreender e justificar.

No entanto, mesmo tendo em conta estes inconvenientes, a adopção do método de classificação das Árvores de Decisão justifica-se pelo facto de ser o que nos permite analisar exaustivamente a interligação das variáveis, o que é essencial para a análise do problema proposto, que consiste na análise do comportamento do consumidor face à montra, no sentido de averiguar que comportamento se concretiza na entrada na loja e na compra. Na análise do comportamento do consumidor é fulcral saber qual o valor assumido para cada uma das variáveis que fazem parte da base de dados para se poder aferir quanto à importância da montra na captação de potenciais clientes. De facto, é tão relevante classificar se um indivíduo/grupo entra, como saber que comportamento teve em relação à montra que esteve na origem dessa classificação, para desta forma analisar em que medida é que a decisão de entrar poderá ter ou não derivado da montra e do seu comportamento face à mesma, e assim se poderem tomar decisões e resoluções que potenciem a montra como um importante instrumento de comunicação da marca e da loja, e de captação de clientes.

Vamos de seguida apresentar como serão aplicadas ao nosso problema, as Árvores de Decisão.

3.3. APLICAÇÃO DAS ÁRVORES DE DECISÃO

No presente trabalho o programa usado para gerar as Árvores de Decisão foi o Weka, que disponibiliza vários métodos de aprendizagem e de Análise de Dados. Para a obtenção das Árvores de Decisão utilizou-se o método J48.

Este método é aplicado a um conjunto de dados, que serão usados para construir o modelo. No entanto, é necessário antes de mais escolher em relação a que atributo se pretende classificar os dados.

Posteriormente é necessário definir o nível de confiança que será usado para podar a árvore. Por definição é usado um nível de confiança de 25%, isto é os atributos que possuam um nível de confiança igual ou superior a 25% são usados como nó da Árvores de Decisão, caso contrário, o nó é substituído por uma folha, ou seja, o atributo é substituído por uma classe, que será a mais frequente em todos os ramos que derivam desse nó. [19]

Para testar a árvore de decisão o programa Weka oferece as seguintes opções de teste [17]: “use training set”, “supplied test set”, “percentage split” e “cross validation”. Na primeira opção os dados de treino são usados para testar o modelo. Na opção “supplied test set” é especificado um conjunto de dados separado que contém os dados para testar o modelo. Na opção “percentage split”, uma percentagem do conjunto de dados de treino do modelo será usada para o testar. Na opção “cross validation” os dados usados para construir o modelo são divididos em n conjuntos, de modo que n-1 conjuntos são usados para construir o modelo e um conjunto para o testar. Este processo será repetido n vezes para que cada um dos n conjuntos de dados seja usado uma vez como conjunto de teste.

Em relação ao output que é gerado, quando estamos perante modelos para problemas com duas classes, como é o caso dos modelos que serão apresentados no capítulo 5, o método gera uma matriz de confusão, que permite analisar a capacidade de previsão do modelo indicando os casos previstos correctamente e incorrectamente [20].

Como podemos analisar na Tabela 1 a matriz de confusão dá a informação dos casos classificados correctamente como sendo da classe (a) (“Verdadeiros Positivos”), dos casos classificados correctamente como sendo da classe (b) (“Verdadeiros Negativos”), dos casos classificados erradamente como sendo da classe (b) mas que são da classe (a) (“Falsos Negativos”) e dos casos classificados erradamente como sendo da classe (a) mas que são da classe (b) (“Falsos Positivos”).

A	B	
(VP) “Verdadeiros Positivos”	(FN) “Falsos Negativos”	a
(FP) “Falsos Positivos”	(VN) “Verdadeiros Negativos”	b

Tabela 1: Diagrama de uma matriz de confusão

A matriz de confusão obtida através do método J48 usando o modelo das Árvores de Decisão, permite analisar a exactidão de cada classe, fornecendo os valores para cada uma, da taxa dos “Verdadeiros Positivos” e dos “Falsos Positivos” e para as medidas Precision Recall e F-Measure [20].

A taxa dos “Verdadeiros Positivos” é dada pelo número de exemplos classificados correctamente em cada classe sobre o total de exemplos dessa classe, ou seja:

$$\text{Taxa dos “Verdadeiros Positivos” para a classe (a)} = (VP) / [(VP)+(FN)]$$

$$\text{Taxa dos “Verdadeiros Positivos” para a classe (b)} = (VN) / [(FP)+(VN)]$$

A taxa dos “Falsos Positivos” é dada pelo número de casos de uma classe classificados incorrectamente no total de casos classificados como sendo dessa classe:

$$\text{Taxa dos “Falsos Positivos” para a classe (a)} = (FP) / [(FP)+(VN)]$$

$$\text{Taxa dos “Falsos Positivos” para a classe (b)} = (FN) / [(VP)+(FN)]$$

A medida Recall corresponde à taxa dos “Verdadeiros Positivos”. Avalia o número total de previsões feitas correctamente face ao número ideal de previsões que deveriam ter sido feitas.

A medida Precision equivale à percentagem de casos classificados correctamente numa dada classe, sobre a totalidade de casos classificados como sendo dessa classe. Assim:

$$\text{Precision para a classe (a)} = (VP) / [(VP)+(FP)]$$

$$\text{Precision para a classe (b)} = (VN) / [(FN)+(VN)]$$

Esta medida corresponde assim às previsões feitas correctamente, em relação ao número total de previsões feitas para uma classe.

A F-Measure combina as medidas Recall e Precision:

$$\frac{2 * \text{recall} * \text{precision}}{\text{recall} + \text{precision}}$$

Esta medida será tanto maior quanto maiores os valores obtidos para as medidas recall e precision, ou seja, quanto melhor for o modelo de classificação obtido, uma vez que tal significa menor número de exemplos classificados incorrectamente.

Após se ter analisado as questões que se colocam à relação entre o consumidor e a marca, e o papel da montra; e após a análise dos problemas relacionados com a análise de dados e dos métodos ao nosso alcance para a realizar, vamos passar à análise dos dados em termos exploratórios e aplicar o método de classificação das Árvores de Decisão. Pretendemos assim ter um melhor conhecimento dos dados em causa e responder às questões colocadas na secção 1.2.

4. ANÁLISE EXPLORATÓRIA DE DADOS

Este capítulo tem por objectivo a apresentação das variáveis, bem como a apresentação do método de recolha dos dados que estão a ser objecto de estudo e a realização de uma análise exploratória dos mesmos. Para a realização desta análise exploratória, numa primeira fase, efectuou-se uma análise preliminar dos dados, que permitiu detectar algumas incoerências que foram devidamente corrigidas e referidas. Após estas correcções, a análise de algumas estatísticas descritivas e de algumas medidas de localização, como a mediana e a análise por quartis, permitirá confirmar a coerência dos dados apresentados. Após este estudo será apresentada uma análise descritiva das variáveis, com o objectivo de compreender o comportamento das mesmas.

4.1. APRESENTAÇÃO DAS VARIÁVEIS

As variáveis utilizadas para realizar este estudo bem como os valores que cada uma assume são apresentados de forma sucinta na Tabela 2. Como se pode verificar as variáveis podem ser divididas em quatro grupos.

- O primeiro permite identificar o dia em que ocorreu a recolha de dados, isto é, se os dados foram recolhidos na época de saldos ou da nova colecção;
- O segundo grupo de variáveis permite caracterizar a deslocação no corredor, isto é, se o indivíduo/grupo observado vem da direita ou da esquerda, se desloca-se no interior ou no exterior do corredor. É importante referir que quando se refere direita ou esquerda e/ou interior ou exterior do corredor significa sempre a posição do indivíduo transeunte em relação à montra.
- Um terceiro grupo de variáveis permite caracterizar o comportamento do indivíduo/grupo face à loja.
- O último grupo de variáveis permitem caracterizar os indivíduos/grupos observados.

Grupo das Variáveis	Variáveis	Categorias	
Dia da Gravação	Dia da Gravação	1	Dia 9 de Setembro de 2006
		2	Dia 16 de Setembro de 2006
Caracterização da deslocação no corredor	Comportamento no corredor	1	Vem da direita
		2	Vem da esquerda
	Deslocação no interior / exterior do corredor	1	Vem no lado da montra
		2	Vem no lado oposto à montra
Caracterização do comportamento do indivíduo face à montra.	Não passa/passa em frente da montra	0	Não Passa
		1	Passa
	Não olha/olha para a montra	0	Não Olha
		1	Olha
	“Ataque” à (para ver a) montra	0	Sem Ataque
		1	Ataque natural
		2	Ataque de aproximação
	Não pára ou pára	0	Não pára
		1	Pára
	Movimento em frente da montra	0	Sem movimento
		1	Quieto
		2	Há movimento
Caracterização do comportamento face à loja	Comportamento em relação à montra	0	Sem comportamento
		1	Fixo à Direita
		2	Fixo ao Centro
		3	Fixo à Esquerda
	Observação da montra	4	Combinação: 1 e/ou 2 e/ou 3
		0	Sem Comportamento
	Hora Início Observação da Montra	1	Foca
		2	Varre
	Hora Fim Observação da Montra	---	Tempo que o indivíduo dispense a observar a montra
		---	Tempo em segundos
Caracterização do comportamento face à loja	Entra ou não entra (na loja)	0	Não entra
		1	Entra
	Hora Entrada na Loja	---	Tempo que o indivíduo dispense no interior da loja
		---	Tempo em segundos
	Tempo de permanência na loja	0	0 Sacos
		1	1 Saco
		2	Mais do que um Saco
	Compras (tipo de compra)	0	não compra
		1	Passa na montra, olha e pára: compra
		2	Passa na montra, olha e não pára: compra
		3	Passa na montra, não olha: compra
Caracterização dos consumidores.	Compras (quem compra)	4	Não passa na montra: compra
		0	Ausência de actividade
		1	Homem Criança
		2	Homem Jovem
		3	Homem Adulto
		4	Homem Terceira - Idade
		5	Mulher Criança
		6	Mulher Jovem
		7	Mulher Adulta
		8	Mulher Terceira - Idade
		9	Grupo
Caracterização dos consumidores.	Indivíduos do grupo (n.º indivíduos).	---	N.º Indivíduos que olham/entram.
	Sexo (de cada indivíduo do grupo)	1	Feminino
		2	Masculino
	Idade (de cada indivíduo do grupo)	1	Criança
		2	Jovem
		3	Adulto
		4	Terceira - Idade

Tabela 2: Apresentação das Variáveis

De seguida é apresentada uma breve descrição de cada variável [21]:

- *Dia da gravação*: Representa os dias em que foram registadas em vídeo as imagens que serviram para o “estudo de caso”. Permite assim identificar a amostra. Se esta variável assume o valor 1 estamos perante as observações recolhidas durante a época de saldos, no dia 9 de Setembro de 2006; se assume o valor 2 estamos perante observações recolhidas no dia 16 de Setembro de 2006, já com nova colecção.
- *Comportamento no corredor (deslocação)*: Indica se o indivíduo vem da direita ou esquerda e significa sempre a posição do indivíduo transeunte em relação à montra.
- *Deslocação no interior / exterior do corredor*: Pretende identificar se o indivíduo se desloca no lado da montra ou no lado oposto à montra, sendo que mais uma vez significa a posição do indivíduo transeunte em relação à montra.
- *Não passa/passa em frente da montra*: Indica se os indivíduos que se deslocam pelo corredor passam em frente da montra ou não.
- *Não olha / olha para a montra*: Divide os transeuntes em duas classes, aqueles que olham para a montra e os que não olham.
- *“Ataque” à (para ver a) montra*: Entenda-se por “ataque” o comportamento do transeunte ao desviar-se da posição por onde segue até à montra, para a observar, isto é, existe uma necessidade de desvio da sua trajectória “natural” para observar a montra pelo que há assim um estímulo evidente da montra sobre o transeunte. Assim “Ataque Natural” significa que o indivíduo vê a montra mas não se desvia da sua rota; “Ataque de Aproximação” que se desloca para a montra; “Sem Ataque” destina-se a catalogar os indivíduos que não olham para a montra.

- *Não pára / pára*: Diferencia os que param dos que não param na montra.

• *Movimento em frente da montra*: Sem movimento significa que o transeunte não passa em frente da montra. Como todas as variáveis têm de registar todos os indivíduos da amostra, esta estratégia permitiu indicar aqueles transeuntes que mesmo deslocando-se no corredor não passam em frente da montra mas que ainda assim entram na loja. Esta variável permite-nos saber se o transeunte assume em frente da montra, uma posição de “estátua” (pára) ou conjugando com a variável “Ataque à (p/ ver) montra” se faz pequenas deslocações com o objectivo de a observar ou ainda se observou a montra à medida que se ia deslocando sem parar.

• *Comportamento em relação à montra*: Sem comportamento, significa que o transeunte não passou em frente da montra. Esta variável permite saber em qual das partes da montra o observador preferiu focar a sua atenção, isto é, se fixou o seu olhar no lado direito, lado esquerdo ou no centro da montra ou ainda se preferiu ver a montra como um todo deslocando-se ao “longo” e observando-a.

• *Observação da montra*: Entenda-se por “focar” olhar para uma determinada posição da montra; “varrer”, olhar para a montra no seu todo. No primeiro caso o indivíduo observa um ponto da montra, já no segundo caso o indivíduo observa vários pontos da montra. É importante referir que um indivíduo para “focar” (“prender” a atenção num ponto da montra – direita, centro ou esquerda) começa por “varrer” a montra e, neste acto contínuo, observa um ponto que o atrai mais fixando assim a sua atenção nesse ponto, permitindo dizer que houve um ponto focal ou lado da montra que prendeu mais sua a atenção.

• *Tempo de paragem ou observação (se não pára) da montra*: Regista o tempo de observação da montra, tenha o indivíduo parado ou observado a montra em andamento.

- *Entra ou não entra (na loja):* Diferencia os indivíduos, entre aqueles que entram daqueles que não entram na loja.

- *Tempo de permanência na loja:* Regista em segundos o tempo de permanência do indivíduo na loja. Assim, quando o tempo de permanência na loja for zero é obvio que o transeunte não entrou.

- *Compras (n.º sacos):* Regista o n.º de compras através da observação do n.º de sacos. Observando o indivíduo à entrada e à saída da loja podemos registar o diferencial do n.º de sacos permitindo saber se o indivíduo efectuou ou não compras.

- *Compras (tipo de compra):* Conjugando esta variável com todas as outras anteriores podemos tentar induzir sobre o tipo de compra que o sujeito efectuou. Temos assim que quando o sujeito “não compra”, pode ter entrado e não comprado ou não ter comprado porque não entrou na loja; “passa na montra – olha e pára: compra”, sabemos que o sujeito foi motivado pela montra a parar, observou-a, entrou na loja e comprou; “passa na montra – olha e não pára: compra”, olhou para a montra não parou, mas podemos saber se houve algum interesse por parte do transeunte em observar a montra, mesmo não parando, através do recurso à variável tempo de observação da montra; “passa na montra – não olha: compra”, a montra não teve qualquer efeito sobre a compra pois o indivíduo não a observou; “não passa na montra: compra”, este sujeito entrou directamente na loja, sendo a compra completamente “limpa” da influência da montra.

- *Compras (quem compra):* Esta variável diz quem efectuou a compra. Divide as pessoas em quatro categorias: criança, jovem, adulto e terceira idade referindo ainda se cada uma destas categorias é homem ou mulher.

- *Indivíduos do grupo (n.º indivíduos):* Esta variável regista o número de indivíduos que formam o grupo observado. Os grupos em estudo vão desde o

“grupo” singular, constituído por uma pessoa, até ao limite máximo que foi observado de sete pessoas.

- *Sexo (de cada indivíduo do grupo)*: Estas variáveis registam cada indivíduo do grupo pelo seu sexo, sendo definido assim se é do sexo feminino ou do sexo masculino.

- *Idade (de cada indivíduo do grupo)*: Na impossibilidade de se saber a idade de cada indivíduo da amostra, a idade cronológica de cada indivíduo do grupo é dividida em quatro escalões: criança, jovem, adulto e indivíduo da terceira idade.

Apresentada a base de dados objecto deste estudo, interessa agora saber como foram obtidos estes dados de caracterização da observação da montra da loja Salsa no Norte Shopping.

4.2. RECOLHA DE DADOS

Os dados apresentados anteriormente foram recolhidos através de duas câmaras de filmar estrategicamente dispostas, uma no interior da montra, que registava a deslocação dos transeuntes no corredor do shopping, e a sua empatia pela montra; outra dentro da loja, permitindo observar se efectuavam compras.

As observações foram recolhidas na Loja “Salsa” no Norte Shopping nos dias 9 e 16 de Setembro de 2006. A primeira amostra foi recolhida na época dos saldos e a segunda com a nova colecção Outono/Inverno. Após a recolha das imagens, estas foram observadas e registadas, por António Ribeiro, no âmbito da sua dissertação em marketing [21] o que permitiu a construção da base de dados em causa.

Com este método de recolha de dados pretendeu-se captar o comportamento do consumidor, sem que este fosse influenciado pela consciência de que estavam a ser analisados, apesar de se ter presente que esta forma de recolha de dados, embora isenta

desse tipo de enviesamento, é mais susceptível de gerar erros de registo das imagens observadas, ou mesmo dificuldade, face à imprecisão de dada imagem observada [22].

Neste sentido, revelou-se importante a realização de uma análise preliminar dos dados que nos permitisse concluir acerca da coerência dos mesmos, o que é descrito na secção posterior.

4.3. ANÁLISE PRELIMINAR

Antes de se iniciar o estudo em causa, interessa averiguar acerca da coerência dos dados apresentados, isto é, se o número de linhas está correcto, se existem linhas repetidas, se as variáveis têm o formato correcto, se o valor máximo e mínimo de cada variável está correcto, se os dados em falta estão correctamente identificados, enfim detectar eventuais erros que possam por em causa análise que se vai realizar.

Realizada esta análise detectaram-se as seguintes incorrecções, e procedeu-se à respectiva rectificação:

- Observação 2470: Tempo de permanência na loja negativo. Procedeu-se à alteração da hora de saída de 13:03 para 14:03. Consequentemente tempo de permanência na loja passou para 534 segundos;
- Observação 3322: Tempo de permanência na loja negativo, pelo que se alterou a hora de saída de 18:03 para 19:03. O tempo de permanência na loja passou para 1073 segundos;
- Observações 75, 113 e 3872: A variável “Entra Vs Não Entra” na loja assume o valor 1, ou seja, “entra”. No entanto o tempo de permanência na loja é 0 segundos logo esta variável tem de ser “não entra”, ou seja, 0.
- Observação 216: A variável “Sexo do indivíduo 1” assume o valor 3, porém esta variável só pode assumir o valor 1, caso o indivíduo seja do sexo Feminino ou 2 caso seja do sexo Masculino. Assim o sexo foi alterado para 2, ou seja, Masculino.
- Observação 4026: As variáveis “Sexo do indivíduo 1” e “Sexo do indivíduo 2” assumem o valor 3, porém estas variáveis só podem assumir o valor 1, caso o indivíduo seja do sexo Feminino, ou 2 caso seja do sexo Masculino; e no caso da variável “Sexo do

indivíduo 2” o valor 0, caso não exista um segundo indivíduo. Assim o sexo de ambas as variáveis foi alterado para 2, ou seja, Masculino;

- Observações 92, 103, 1961, 2636: O indivíduo passa em frente da montra, olha para a montra, não pára e compra, logo o valor da variável “Tipo de Compra” é 2 e não 1.

De seguida, na Tabela 3 e na Tabela 4, é apresentado o valor de algumas estatísticas descritivas das variáveis em estudo.

Variáveis Qualitativas	N	Mínimo	Máximo
DiaGravacao	4049	1	2
ComportamentoCorredor	4049	1	2
DeslocacaoCorredor	4049	1	2
NaoPassaVsPassaFrenteMontra	4049	0	1
NaoOlhaVsOlhaMontra	4049	0	1
AtaqueMontra	4049	0	2
NaoPáraVsPára	4049	0	1
MovimentoMontra	4049	0	2
ComportamentoMontra	4049	0	4
ObservacaoMontra	4049	0	2
EntraVsNaoEntra	4049	0	1
TipoCompra	4049	0	4
QuemCompra	4049	0	9
Sexo1	4049	1	2
Sexo2	4049	0	2
Sexo3	4049	0	2
Sexo4	4049	0	2
Sexo5	4049	0	2
Sexo6	4049	0	2
Sexo7	4049	0	2
Idade1	4049	1	4
Idade2	4049	0	4
Idade3	4049	0	4
Idade4	4049	0	4
Idade5	4049	0	4
Idade6	4049	0	4
Idade7	4049	0	4

Tabela 3: Estatísticas Descritivas das Variáveis Qualitativas

Variáveis Quantitativas	N	Mínimo	Máximo	Média	Desvio Padrão
TempoObservacao(Seg)	4049	0	245	3.41	7.19
TempoLoja(seg)	4049	0	3786	204.97	424.12
NumSacosCompras	4049	0	2	0.08	0.28
NúmeroIndivíduos	4049	1	7	1.96	0.90
HoraInicioObservação	4049	100146	270750	----	----
HoraFimObservação	4049	100146	270752	----	----
HoraEntradaLoja	4049	0	235718	----	----
HoraSaídaLoja	4049	0	235748	----	----

Tabela 4: Estatísticas Descritivas das Variáveis Quantitativas

Os valores apresentados permitem verificar a coerência da base de dados objecto de estudo, o que foi confirmado com a análise das medidas de localização, que são também um importante instrumento para averiguar da coerência dos dados apresentados, nomeadamente a média, a mediana e a análise por quartis.

A média é apresentada na Tabela 4, e possui os valores esperados. Assim podemos, por exemplo, concluir que em média cada indivíduo/grupo permanece 204,97 segundos na loja, e que em média cada observação é constituída por grupos de 2 indivíduos. A mediana é uma medida mais resistente do que a média a “outliers”, sendo os seus valores apresentados na Tabela 5, verificando-se que são coerentes em todas as variáveis.

Variável	Mediana
DiaGravacao	2
ComportamentoCorredor	1
DeslocacaoCorredor	2
NaoPassaVsPassaFrenteMontra	1
NaoOlhaVsOlhaMontra	1
AtaqueMontra	1
NaoPáraVsPára	0
MovimentoMontra	1
ComportamentoMontra	2
ObservacaoMontra	2
TempoObservacao(Seg)	2
EntraVsNaoEntra	0
TempoLoja(seg)	0
NumSacosCompras	0
TipoCompra	0
QuemCompra	0
NúmeroIndivíduos	2
Sexo1	1
Sexo2	1

Variável	Mediana
Sexo3	0
Sexo4	0
Sexo5	0
Sexo6	0
Sexo7	0
Idade1	2
Idade2	2
Idade3	0
Idade4	0
Idade5	0
Idade6	0
Idade7	0

Tabela 5: Mediana das Variáveis em Estudo

Finalmente para a análise dos quartis recorreremos à representação gráfica, através da caixa de extremos e quartis (box-plot ou caixa de bigodes).

Para a análise dos gráficos apresentados de seguida, é importante saber o seguinte [23]:

- Cada box-plot estende-se do 1º (Q_1) ao 3º (Q_3) quartil que correspondem, respectivamente, à base inferior e superior do rectângulo e que representa 50% das observações totais;
- A mediana é representada por uma linha grossa dentro da caixa; se a amostra for aproximadamente simétrica, a linha que corresponde à mediana divide a caixa em duas partes aproximadamente iguais;
- Um outlier, ou melhor um candidato a outlier, é um valor que se situa fora das chamadas barreiras de outliers e que são determinadas como sendo:

$$(Q_1 - 1,5 (Q_3 - Q_1) ; Q_3 + 1,5 (Q_3 - Q_1))$$

(Outlier Inferior, Outlier Superior).

Assim uma observação é considerada outlier moderado (assinalado nos gráficos com um círculo) se estiver entre $Q_1 - 3 (Q_3 - Q_1)$ e $Q_1 - 1,5 (Q_3 - Q_1)$ e entre $Q_3 + 1,5 (Q_3 - Q_1)$ e $Q_3 + 3 (Q_3 - Q_1)$. Será considerado um outlier severo se for inferior a $Q_1 - 3 (Q_3 - Q_1)$ ou superior a $Q_3 + 3 (Q_3 - Q_1)$ (assinalado nos gráficos com um asterisco). Vamos de seguida apresentar os gráficos em causa.

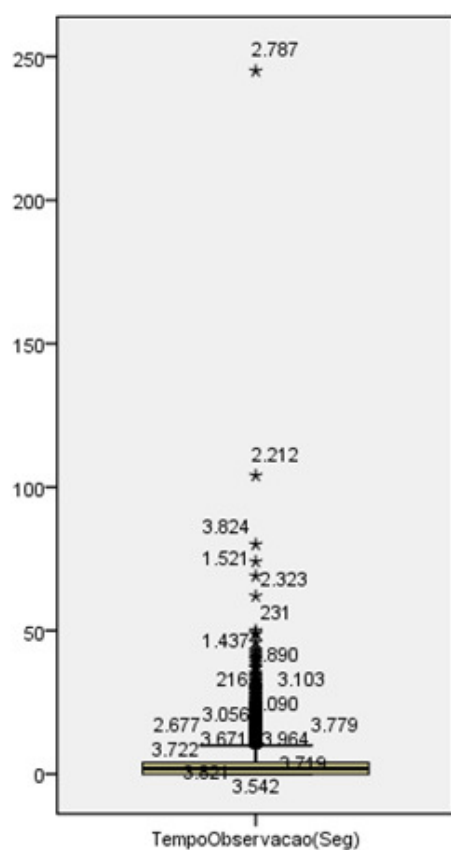


Gráfico n.º 1: Box Plot do Tempo de Observação da Montra

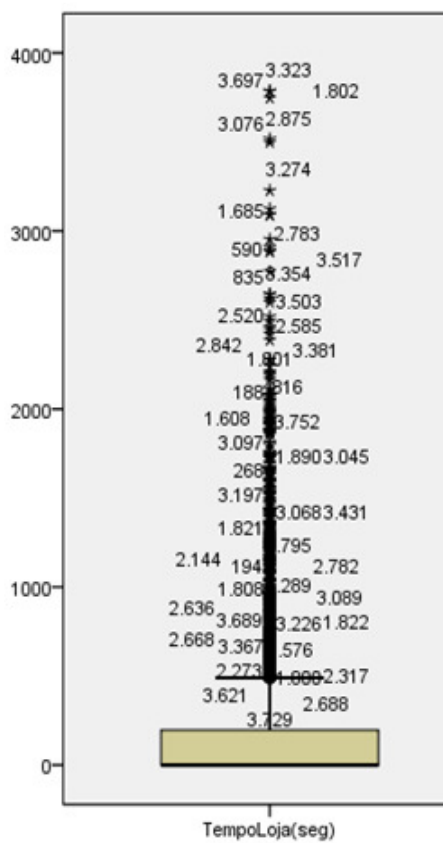


Gráfico n.º 2: Box Plot do Tempo na Loja

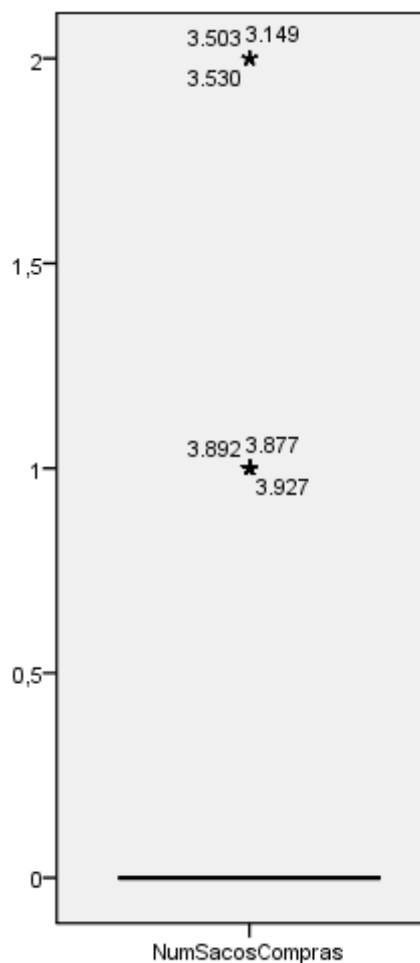


Gráfico n.º 3: Box Plot do Número de Sacos de Compras

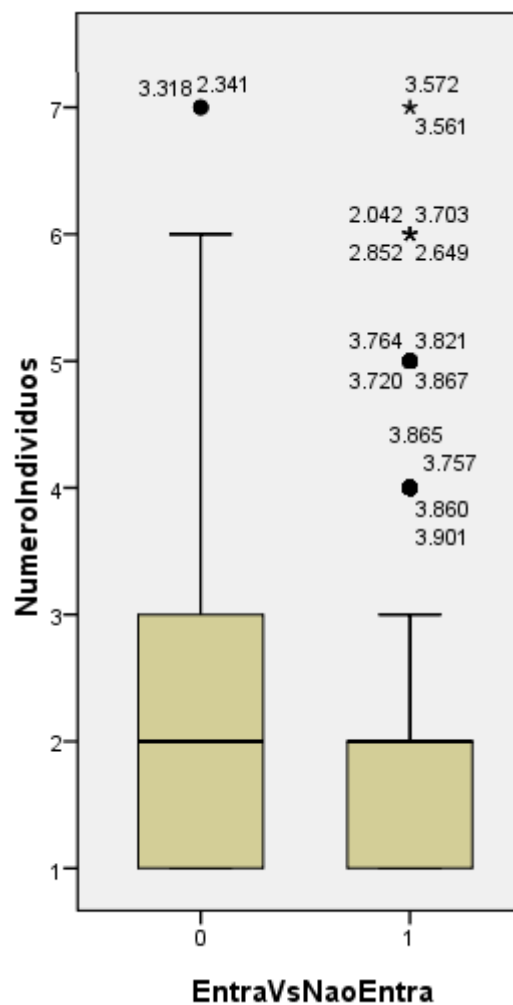


Gráfico n.º 4 Box Plot do Número de Indivíduos que compõe cada Observação

Assim podemos verificar que as variáveis “Tempo na Loja” e “Número de Indivíduos” não são variáveis simétricas, e que todas apresentam outliers superiores, moderados e severos. Na variável “Tempo de Observação da montra” há uma observação que se destaca (outlier severo), a observação 2787, na qual um indivíduo observa a montra durante cerca de quatro minutos. Em relação ao número de sacos de compras, apenas em três observações se registam mais do que um saco de compras. Finalmente em relação ao número de indivíduos que não entra na loja, apenas duas observações eram compostas por 7 indivíduos e em relação ao número de indivíduos que entra na loja, apenas 12 observações eram compostas por quatro ou mais indivíduos.

Após esta primeira abordagem às variáveis em estudo, interessa agora saber qual o comportamento de cada uma delas, e de que forma se relacionam, de modo a se conhecer

melhor a base de dados em estudo, pelo que de seguida é realizada uma análise descritiva das variáveis.

4.4. ANÁLISE DESCRITIVA DAS VARIÁVEIS

Com o objectivo de se compreender o comportamento das variáveis, para se ter uma perspectiva global dos valores que assumem nas diferentes situações apresentadas, e de como os casos da amostra recolhida se distribuem pelos diferentes atributos de cada classe, é aqui realizada uma análise descritiva das variáveis.

Assim, numa primeira fase, vai-se analisar o comportamento de cada variável e será definido o perfil do indivíduo/grupo que entra na loja e que efectivamente compra.

Dado que a loja objecto de estudo está inserida num centro comercial, é também analisada a afluência dos indivíduos/grupos ao shopping de modo a se descobrir qual a altura do dia que regista maior número de observações, maior número de entradas na loja e maior número de compras.

Por fim, interessa avaliar o impacto da montra na decisão de entrar ou não na loja, isto é, analisar em que medida o indivíduo entra ou não na loja em função do comportamento no corredor; da deslocação no corredor; do facto de passar ou não em frente da montra; de olhar ou não para a montra; do ataque à montra; do facto de parar ou não em frente da montra; do movimento em frente à montra; do comportamento face à montra; e da observação da montra.

4.4.1. ANÁLISE DO COMPORTAMENTO DAS VARIÁVEIS

Nesta secção vai-se analisar o comportamento das variáveis que estão a ser objecto de estudo e análise. Assim analisar-se-á como as observações se repartem pelas diferentes

variáveis e em que medida é possível estabelecer relações entre elas. Da mesma forma pretende-se caracterizar o perfil do indivíduo/grupo que entra na loja e que efectivamente compra.

4.4.1.1. EQUILÍBRIO DA AMOSTRA

Numa primeira análise verifica-se que estamos perante uma amostra equilibrada, isto é uma amostra que tem um número aproximado de casos em cada um dos dias em que foram recolhidas as observações (49,35% e 50,65%, no dia 9 e 16 de Setembro respectivamente). Outra questão relevante é saber como os casos se distribuem pela variável “Entra ou Não na Loja” e “Compra ou Não”, pois é para estas variáveis que pretendemos obter regras de classificação.

Assim verifica-se que cerca de 50% do total de indivíduos observados entra na loja, embora se verifique que entram mais pessoas na loja no dia 16 de Setembro, como se pode constatar no Gráfico n.º 5 (cerca de 55,5% dos indivíduos observados nesse dia face a 44% dos indivíduos que entram na loja no dia 9). Isto pode ser fruto do facto de o dia 16 de Setembro ser o dia de início da nova colecção e tal suscitar maior curiosidade e interesse nos consumidores.

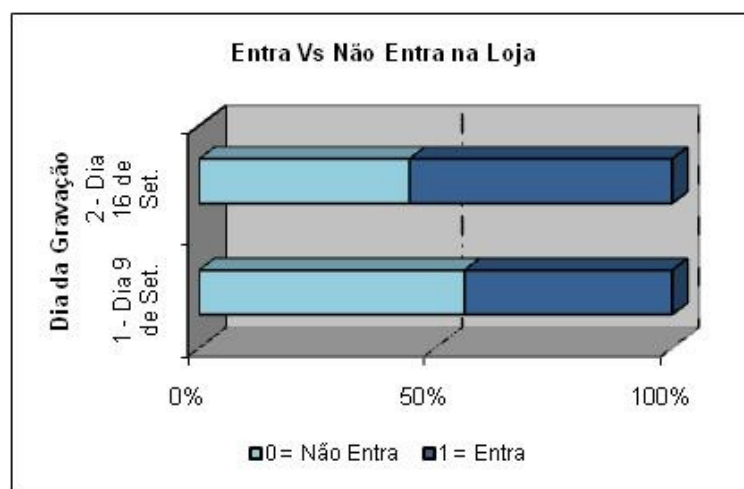


Gráfico n.º 5: Distribuição dos dados pela variável “Entra ou Não Entra na Loja”

Por outro lado a grande maioria dos indivíduos/grupos observados não realizam compras, o que sucede nos dois dias observados, como se demonstra no Gráfico n.º 6.

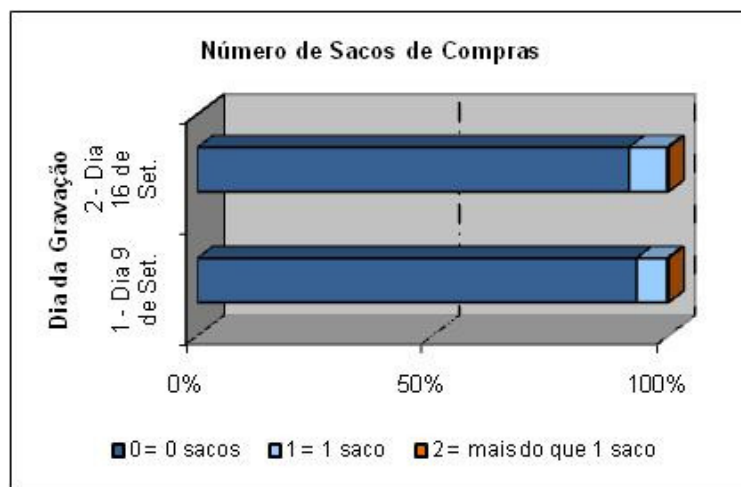


Gráfico n.º 6: Distribuição dos dados pela variável “Número de Sacos de Compras”

4.4.1.2. DESLOCAÇÃO NO CORREDOR

Uma outra abordagem que é interessante realizar é descrever a deslocação do indivíduo/grupo no corredor. Constata-se que cerca de 67% dos indivíduos observados vêm da direita da montra, isto é, podem entrar na loja sem passar pela montra, enquanto 33% dos indivíduos vem da esquerda, isto é, se tiverem intenção de entrar na loja passam obrigatoriamente na montra. Por outro lado, cerca de 47% dos indivíduos desloca-se no interior do corredor, isto é próximo da montra, enquanto 53% desloca-se no lado oposto da montra. De facto, é interessante verificar que apesar de nos dois dias, os indivíduos que passam no corredor em frente da loja virem maioritariamente da direita, a maioria dos indivíduos deslocam-se no lado oposto à montra, o que não seria de esperar tendo em conta as regras de circulação que são intrínsecas à nossa cultura.

4.4.1.3. COMPORTAMENTO FACE À MONTRA

Outra análise relevante para o estudo que se pretende realizar é a caracterização do comportamento dos indivíduos/grupos observados face à montra. Verificou-se que a grande maioria dos indivíduos que passa em frente da montra, olha efectivamente para a montra, nos dois dias. De facto, cerca de 86% dos indivíduos que se deslocam no interior do corredor, isto é, junto à montra, passam efectivamente em frente dela, enquanto 77% dos indivíduos que se deslocam no lado exterior do corredor, também passam em frente da montra, como é visível no Gráfico n.º 7.

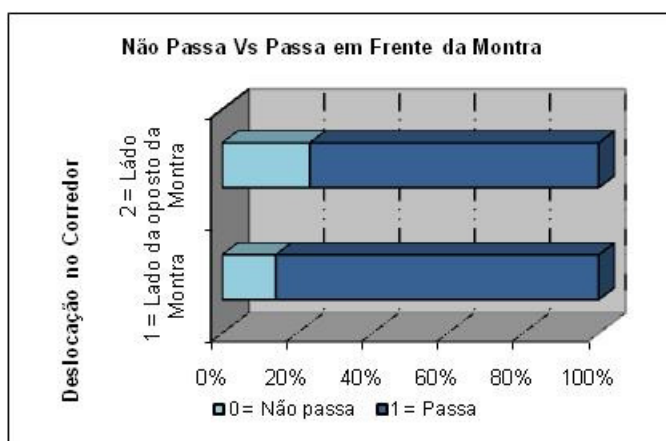


Gráfico n.º 7: Distribuição dos dados pelas variáveis “Passa ou Não Passa em frente da montra” e “Deslocação no interior/exterior do corredor”.

Nos dois dias, mais de 50% dos indivíduos tem um ataque natural face à montra, isto é, não se desviam da sua rota para ver a montra. É ainda importante referir que a percentagem de indivíduos que não faz qualquer ataque à montra, isto é, nem sequer olha para ela, é significativamente maior do que os que realizam um ataque de aproximação, isto é, dos que têm uma necessidade de desvio da sua trajectória “natural” para observar a montra e que, assim, revelam um estímulo evidente da montra sobre eles. Daqueles que se desviam da sua trajectória para observar a montra, todos se deslocam no exterior do corredor, exceptuando um único caso (Gráfico n.º 8). Estes manifestam um interesse bastante significativo pela montra.

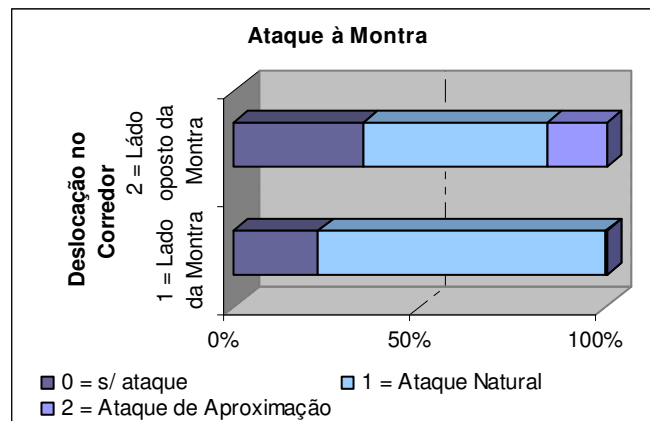


Gráfico n.º 8: Distribuição dos dados pelas variáveis “Ataque à montra” e “Deslocação no Corredor”

A maioria dos indivíduos/grupos não pára para ver a montra, nos dois dias observados (cerca de 82% das observações recolhidas em cada dia). Dos indivíduos/grupos que observam a montra, a maioria movimenta-se ao longo da mesma, nos dois dias, e fixam-se em mais do que uma zona da montra (direita e esquerda ou direita e centro ou esquerda e centro), representado cerca de 48% e de 39% das observações do dia 9 e 16 de Setembro, respectivamente. No entanto, dos indivíduos que observam em particular uma parte da montra, verifica-se que a zona centro da montra é a privilegiada, com cerca de 18% e de 21% das observações do dia 9 e 16 de Setembro, respectivamente. Interessa ainda analisar em relação à observação da montra, se é focado um aspecto em particular da montra ou se esta é analisada como um todo (se o indivíduo foca ou varre a montra). Assim verifica-se que 60% olha para a montra como um todo (varre), enquanto 11% foca um determinado aspecto.

Destes 443 indivíduos que focam um dado aspecto da montra, 47% fixa-se à direita, 27% ao centro e 26% à esquerda, consequentemente a parte da montra que capta maior atenção, neste caso, é assim a parte direita da montra. De facto, desta análise verifica-se que quando o indivíduo varre a montra com o olhar, para o fazer, ou fixa-se em mais do que um ponto da montra, ou fixa-se ao centro; mas quando foca um aspecto particular da montra, o maior número de observações verifica-se para os indivíduos que se fixam à direita, seguido do centro e por fim da esquerda. Estas conclusões podem ser visualizadas no Gráfico n.º 9

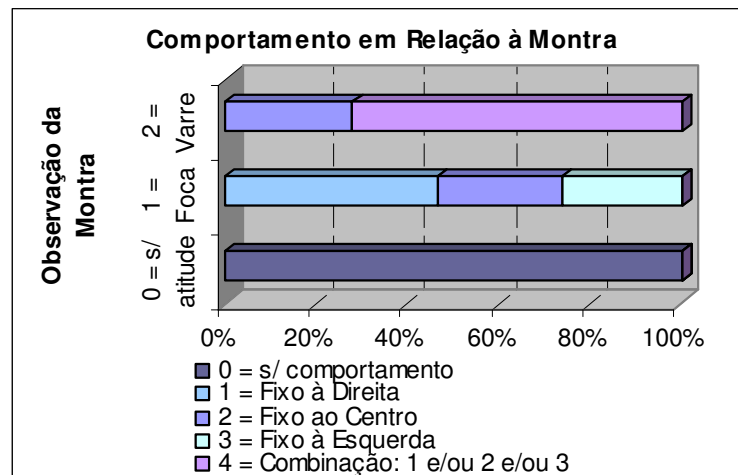


Gráfico n.º 9: Distribuição das observações pelas variáveis “Observação da montra” e “Comportamento em Relação à montra”

De referir que, observando as imagens que foram recolhidas, a zona direita da montra possui um manequim com roupa feminina, a zona esquerda da montra possui mais artigos expostos, pois possuiu dois manequins (um feminino e um masculino), a zona do centro da montra possuiu peças de roupa pousadas, de modo a se posicionarem também a meio da montra em termos de altura, sendo esta a zona da montra mais observada.

A maioria dos indivíduos/grupos que compra, não passa na montra, pelo que compra com base na relação que já possuiu com a marca e com a loja, entrando com intenção ou o objectivo de comprar algo. De facto, apenas 121 indivíduos passam na montra, observam-na e de facto compram.

4.4.1.4. CARACTERIZAÇÃO DOS INDIVÍDUOS /GRUPOS

Analisando os consumidores que são observados, verifica-se que cerca de 32% das observações correspondem a uma única pessoa, enquanto em 68% observações estiveram presentes mais do que uma pessoa: destas 2746 observações, em 1909 estiveram presentes 2 pessoas (cerca de 69%), em 592, 3 pessoas (cerca de 21%), em 191, 4 pessoas (cerca de 7%), 5 pessoas estiveram presente em 39 observações (cerca de

1%), 6 em 10 observações (cerca de 0,4%) e 7 pessoas em 5 observações (cerca de 0,2%). Assim podemos concluir que a maioria das observações foi registada para grupos de duas pessoas. Analisando o número de indivíduos de cada observação com as compras que se registaram, verifica-se que a maioria das compras é também realizada por grupos de duas pessoas, como se pode observar no Gráfico n.º 10.

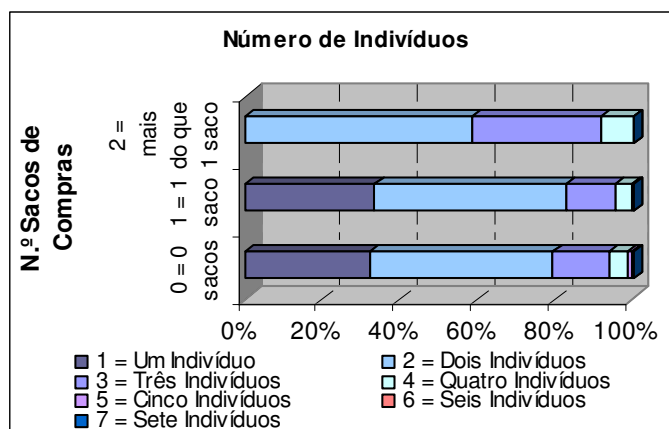


Gráfico n.º 10: Distribuição dos dados pelas variáveis “Número de indivíduos” e “Número de Sacos de Compras”

Podemos ir mais longe na análise dos consumidores observados e definir o perfil do indivíduo/grupo observado, de quem entra na loja e de quem efectivamente compra. Para tal foram criadas cinco novas variáveis:

- Idade do Grupo: Média das Idades dos indivíduos de cada observação;
- Sexo do Grupo: Indica se cada observação é constituída apenas pelo sexo feminino (F), apenas pelo sexo masculino (M), ou por ambos os sexos (Misto);
- % Sexo Feminino: Indica a % de indivíduos na observação que são do sexo feminino;
- % Sexo Masculino: Indica a % de indivíduos na observação que são do sexo masculino;
- “Compra vs Não Compra: será igual a 0 se o indivíduo não compra, isto é se o número de sacos de compras é igual a zero, e será igual a 1 se o número de sacos de compra for igual ou superior a um.

Com base nestas variáveis, numa primeira fase caracterizar-se-á quem observa a montra, isto é, quem entra e quem compra, em termos de facha etária em que se insere, e numa segunda fase, em termos de sexo.

Constata-se assim, como é visível no Gráfico n.º 11, que a maioria dos indivíduos/grupos observados é composto ou por um jovem ou por um grupo de pessoas maioritariamente jovens, sendo estes aqueles que também mais entram na loja, e com maior percentagem no dia 16 de Setembro (no dia 16 cerca de 59% entram na loja, face a 47% no dia 6). De facto, a média das idade do grupo (constituído por uma ou mais pessoas) que efectivamente compra insere-se na categoria “jovem”, nos dois dias observados e existe um maior número de compras no segundo dia de observação, isto é, na época pós saldos, como se pode verificar no Gráfico n.º 12.

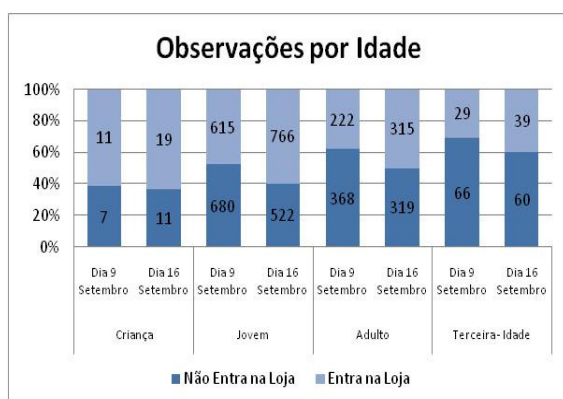


Gráfico n.º 11: Proporção de indivíduos que entra ou não na loja, por idade, nos dois dias observados



Gráfico n.º 12: Número de compras por grupo de idade, nos dois dias de observação

Na amostra predominam os grupos constituídos por ambos os sexos (45,27% das observações), seguido pelos grupos compostos pelo sexo exclusivamente feminino (41,89% das observações). Os grupos constituídos exclusivamente pelo sexo masculino são significativamente inferiores em ambos os dias. Importa ainda referir que os grupos/indivíduos do sexo feminino são os que mais entram na loja face a todos os outros, nos dois dias, como se pode verificar no Gráfico n.º 13. Analisando com maior pormenor os grupos/indivíduos sexo feminino constata-se que a maioria das observações nos dois dias é constituído por um único indivíduo, que entram na loja em maior proporção face aos grupos observados: 475 observações das 828 do dia 6 de Setembro são constituídas por apenas um indivíduo, sendo que destes 261 entram na

loja; e 452 das 868 do dia 19 de Setembro são constituídas por apenas um indivíduo, sendo que destes 286 entram na loja. Analisando quem compra por sexo verifica-se que os grupos que realizam compra são maioritariamente constituídos por pessoas do sexo feminino no dia 9 de Setembro, mas no dia 16, já são maioritariamente constituídos por pessoas de ambos os sexos como se pode constatar no Gráfico n.º 14.

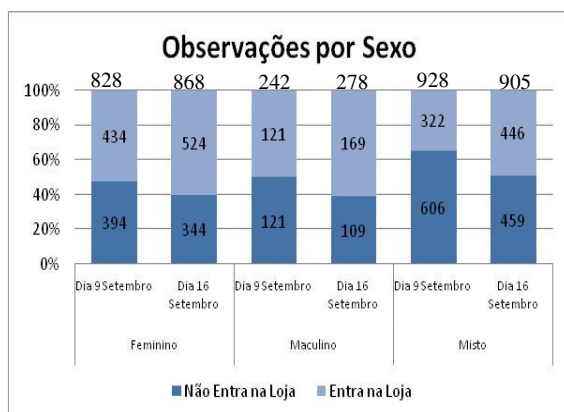


Gráfico n.º 13: Proporção de Observações face à variável Entra ou não na loja, por sexo, por dia de observação

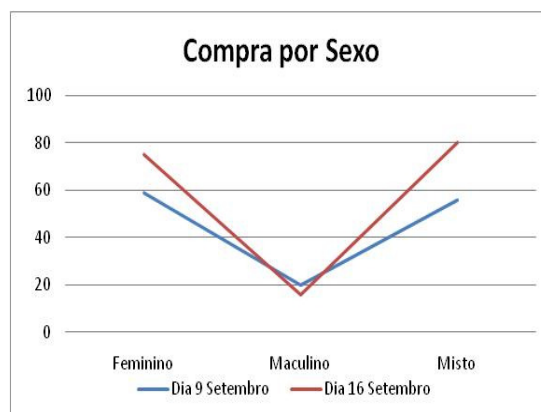


Gráfico n.º 14: Compra por sexo por dia de observação

Os grupos/indivíduos do sexo feminino que compram são maioritariamente constituídos por uma única pessoa (32 das 59 observações do dia 9 de Setembro e 39 das 75 observações do dia 16 de Setembro). Analisando a composição dos grupos constituídos por indivíduos de ambos os sexos consta-se que a maioria é constituído por grupos com 50% de pessoas do sexo feminino e 50% do sexo masculino, sendo estes aqueles que efectivamente mais entram na loja e que mais compram como é visível no Gráfico n.º 15.



Gráfico n.º 15: Composição dos grupos constituídos pelos 2 sexos que realizam compra

Considerando todas as variáveis em conjunto, verifica-se que a amostra é constituída por grupos/indivíduos maioritariamente jovens, e que estes grupos/indivíduos são constituídos maioritariamente por pessoas do sexo feminino ou por pessoas em igual número do sexo masculino e feminino. São estes grupos/indivíduos que efectivamente mais entram na loja. De facto, são os grupos maioritariamente jovens que mais compram em ambos os dias de observação. Estes grupos que realizam a compra fazem-no em maior proporção quando estão presentes indivíduos de ambos os sexos, nos dois dias.

Falta agora responder a duas questões: qual o período de maior afluência dos indivíduos à loja, uma vez que esta está inserida num centro comercial, e descrever o comportamento dos indivíduos/grupos face à montra tendo em conta todas as variáveis de que o descrevem.

4.4.2. AFLUÊNCIA DOS INDIVÍDUOS

Uma vez que a loja objecto de estudo está inserida num centro comercial, será interessante analisar a afluência dos indivíduos/grupos que são observados, isto é, qual será a altura do dia que regista maior número de observações, maior número de entradas na loja e mais compras?

Tendo por base a hora em que cada indivíduo/grupo passa no corredor as observações foram divididas em 6 grupos:

- Manhã: das 10:00h às 12:30h (exclusive)
- Almoço: das 12:30h às 13:00h (exclusive)
- Tarde: das 13:00h às 17:30h (exclusive)
- Final Tarde: das 17:30h às 19:30h (exclusive)
- Jantar: das 19:30 às 21:30h (exclusive)
- Noite: das 21:30 às 24:00h (inclusive)

Para esta análise é importante referir que os dois dias de observação foram num sábado. Os resultados obtidos, apresentados no Gráfico n.º 16, demonstram que o período do dia com maior número de observações é o período da “Tarde” (principalmente no segundo dia de observação), o que deve ser motivado pelo facto de estarmos perante um dia de fim-de-semana. O período definido como a hora de Jantar é nos dois dias o período com menor número de observações. Isto porque independentemente da afluência ao shopping, que provavelmente até será maior nesta altura do dia, poderá ocorrer que as pessoas se concentrem nesta altura na praça da alimentação. Coincidente com esta informação verifica-se que entra mais pessoas na loja no período da tarde, e é nesta altura do dia que há maior número de compras, como é visível no Gráfico n.º 17 e no Gráfico n.º 18, respectivamente.

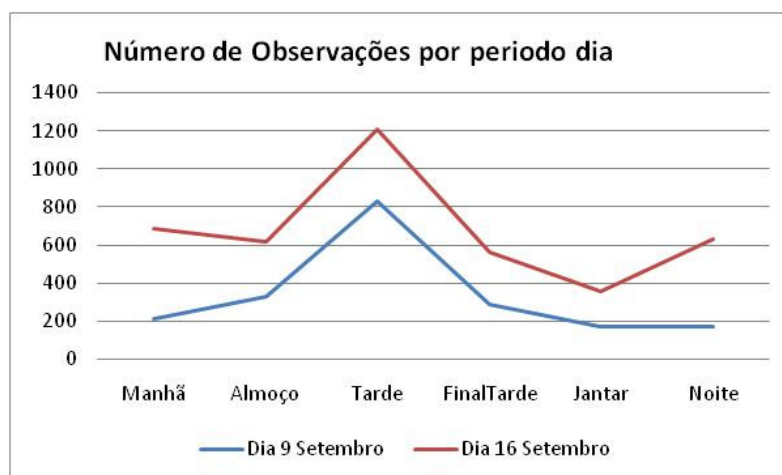


Gráfico n.º 16: Número de observações por período do dia

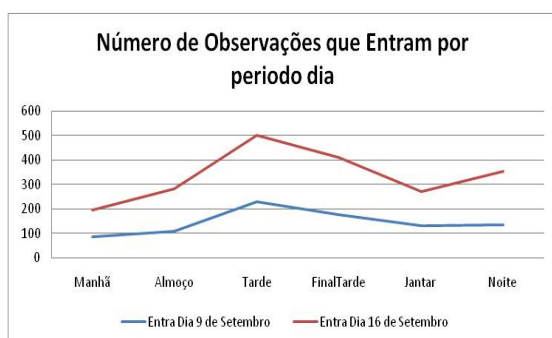


Gráfico n.º 17: Número de Observações em que se entra na loja por período do dia por dia de observação

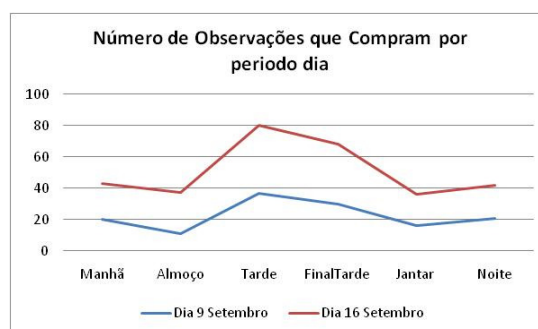


Gráfico n.º 18: Número de observações em que se compra por período do dia por dia de observação

É no entanto importante referir, que a percentagem de compras face ao número de pessoas que entram na loja é maior no período da Manhã e no período do Final da Tarde, nos dois dias: cerca de 22,73% e 16,85% no dia 9 de Setembro e cerca de 21,30% e 16,31% no dia 16 de Setembro.

Concluiu-se assim que existe maior número de observações entre as 13:00 e as 17:00 horas, sendo neste período de tempo que se registam mais entradas na loja e maior número de compras.

Vamos agora descrever o comportamento dos indivíduos/grupos face à montra tendo em conta todas as variáveis de que o descrevem.

4.4.3. DECISÃO DE COMPRA

O objectivo é relacionar as variáveis que descrevem o comportamento do indivíduo face à montra, nomeadamente, “Não passa/passa em frente da montra”, “Não olha/olha para a montra”, “Ataque à (para ver a) montra”, “Não pára ou pára”, “Movimento em frente da montra”, “Comportamento em relação à montra”, “Observação da montra”, “Tempo de paragem ou observação da montra”; com a decisão de compra.

Pretende-se avaliar o impacto da montra na decisão de entrar ou não na loja e analisar em que medida o indivíduo entra ou não na loja em função do comportamento no corredor; da deslocação no corredor; do facto de passar ou não em frente da montra; de olhar ou não para a montra; do ataque à montra; do facto de parar ou não em frente da montra; do movimento em frente à montra; do comportamento face à montra; e da observação da montra. O objectivo é obter algum conhecimento acerca de como as variáveis se relacionam.

Desta análise concluiu-se o seguinte (informação disponível no Anexo n.º 1):

- Em 39 e 70 observações, no dia 9 e 16 de Setembro, respectivamente, houve compra, num universo de 333 e 436 observações em que efectivamente se entrou na

loja, mas que não houve passagem em frente da montra e como tal não houve manifestação de qualquer reacção face à montra. Tal representa cerca de 12% e 16% dos indivíduos que entram na loja no dia 9 e 16 de Setembro, respectivamente.

- Nos restantes casos, o indivíduo passa em frente da montra e:
 - Não manifesta qualquer interesse pela montra, mas entra na loja e em 31 e 45 observações, no dia 9 e 16 de Setembro, houve compra, num universo de 165 e 240 observações, respectivamente, isto é 100% dos indivíduos observados que não passa em frente da montra e não manifesta qualquer interesse por ela, entra na loja, em cada um dos dias, e destes cerca de 19%, em cada um dos dias, realiza compra;
 - Mostra interesse pela montra, isto é olha para a montra, e em 1121 e 912 casos do dia 9 e 16 de Setembro, respectivamente, não entra na loja. Em 379 e 463 observações do dia 9 e 16 de Setembro, o indivíduo entra na loja mas só compra em 65 e 56 casos, respectivamente. Ou seja; cerca de 25% e 34% dos indivíduos do dia 9 e 16 de Setembro respectivamente, que passam em frente da montra e olham para a mesma, entra na loja, mas destes, apenas 17% no dia 9 e 12% no dia 16, efectivamente compra.

Neste capítulo foi possível adquirir um conhecimento dos dados e do comportamento das variáveis em causa, de seguida serão aplicadas técnicas de Data Mining no sentido de se descobrir o perfil de um potencial consumidor, isto é, de quem entra na loja, de quem compra, e de averiguar se esse comportamento é ou não distinto consoante se está na época de saldos ou na época pós-saldos.

5. DATA MINING

O objectivo deste trabalho é analisar o que leva um indivíduo a entrar na loja, o que motiva a compra, e se o comportamento é distinto consoante a época, isto é, se é ou não época de saldos.

Para realizar esta análise foram usadas as variáveis que foram construídas e enunciadas no Capítulo 4, “Compra vs Não Compra”, “Sexo Grupo” e “Média Idade”.

Neste capítulo utilizaremos, para esta análise, as Árvores de Decisão e o método usado para as obter será o J48 do programa Weka.

Este método é aplicado ao conjunto de dados apresentados anteriormente (secção 4.1), obtendo assim modelos de classificação. O primeiro modelo será obtido em relação à variável “Entra ou Não na Loja” e o segundo em relação à variável “Compra ou Não”. Estes serão portanto, os atributos que guardam a classe dos casos observados. Serão assim obtidos quatro modelos de classificação: Entra ou Não Entra na Loja no dia 9 de Setembro; Entra ou Não Entra na Loja no dia 16 de Setembro; Compra ou não no dia 9 de Setembro e Compra ou não no dia 16 de Setembro.

Em relação aos parâmetros usados para obter os modelos será usado um nível de confiança de 25%. Para estimar a taxa de erro dos modelos obtidos será usada a opção de “cross validation” com 10 grupos. O modelo final será criado usando a totalidade dos dados de treino e a taxa de erro estimada deste modelo em novos dados com a mesma distribuição, será a média da taxa de erro obtida nos 10 modelos anteriormente construídos. Esta opção permite usar a quantidade máxima de dados disponível para criar o modelo e ao mesmo tempo obter uma boa estimativa da taxa de erro esperada.

No final deste capítulo dever-se-á conseguir prever, face a um novo caso, se um indivíduo entra ou não na loja e se compra ou não. O objectivo desta análise é estudar o poder de atracção da montra na época de saldos e na época de nova colecção, isto é, em que medida a montra exerce um poder de atracção sobre os indivíduos, podendo assim ser um meio de captação de clientes a explorar.

5.1. QUEM ENTRA

Nesta secção vai-se então analisar quem entra na loja. Para tal não foram consideradas as seguintes variáveis:

- Número de sacos de Compras;
- Tipo de Compra;
- Quem Compra;
- Hora Entrada da Loja
- Hora Saída da Loja
- Tempo na Loja (Seg.)

A exclusão destas variáveis resulta do facto de que todas são consequência de o indivíduo entrar na loja, e portanto o seu valor não estará disponível quando pretendemos prever se o indivíduo irá ou não entrar na loja. Por simplificação, uma vez que se inclui o Tempo de Observação da Montra, não foi considerada nem a Hora de Início da Observação, nem a Hora de Fim da Observação.

Para analisar quem entra na loja vai-se então aplicar o método J48 para se obter Árvores de Decisão em relação à variável “Entra ou Não na Loja”. Isto será feito numa primeira fase para as observações recolhidas no dia 9 de Setembro e numa segunda fase para as observações recolhidas no dia 16 de Setembro.

5.1.1. MODELO DE CLASSIFICAÇÃO OBTIDO PARA A VARIÁVEL “ENTRA OU NÃO NA LOJA” NO DIA 9 DE SETEMBRO

Numa primeira fase, o modelo de classificação foi construído usando 1998 observações que correspondem às recolhidas no dia 9 de Setembro, e é apresentado na Figura n.º 2.

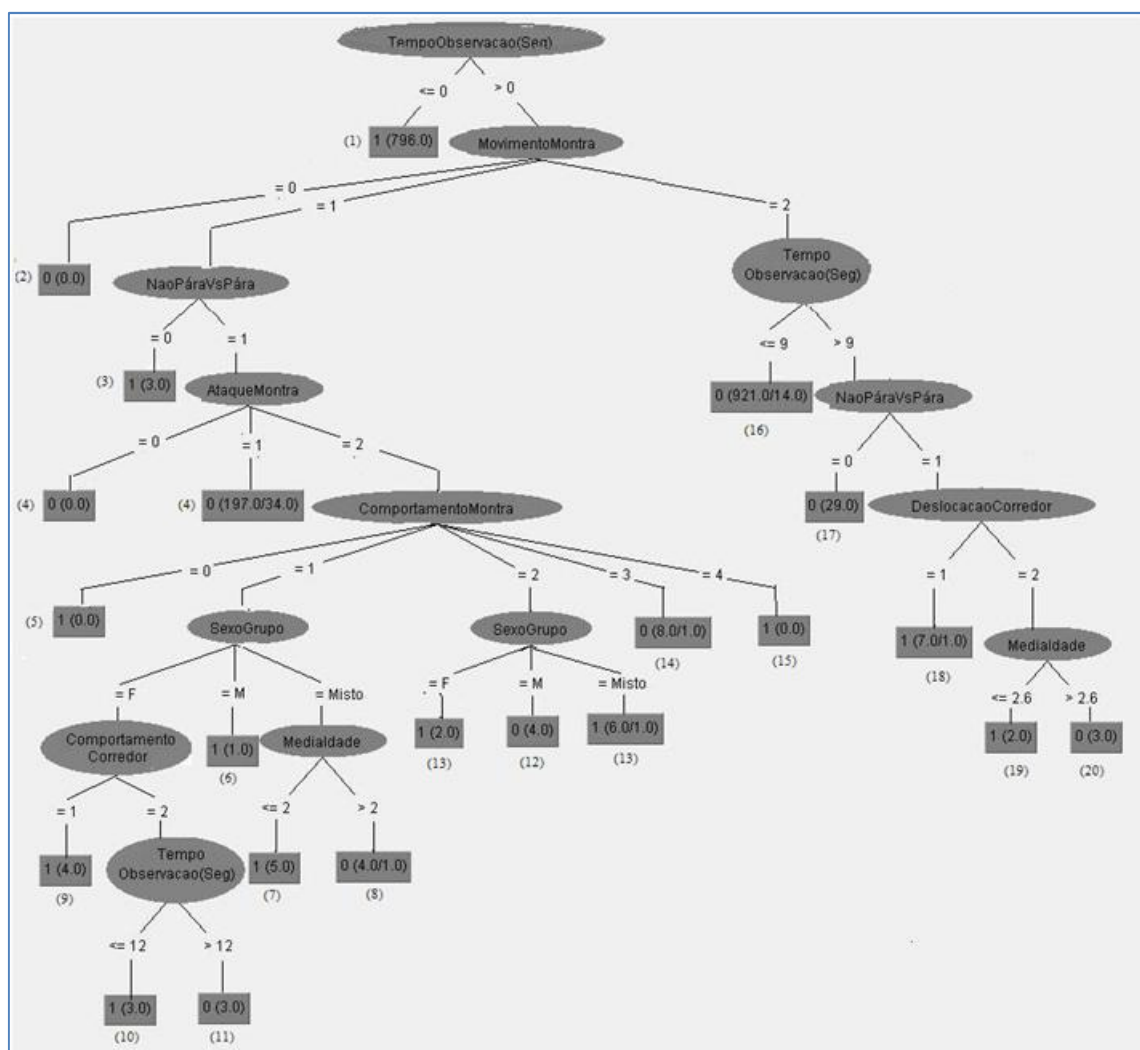


Figura n.º 2: Árvore de Classificação “Entra ou não na loja” no dia 9 de Setembro

Este modelo classifica correctamente 1927 casos como podemos constatar no output gerado pelo modelo J48 do Weka do Anexo n.º 2, resumido na Tabela 6.

		N.º	%
Casos Classificados Correctamente		1927	96,45
Casos Classificados Incorrectamente	Como “Entram na Loja”	10	0,50
	Como “Não Entram na Loja”	61	3,05
	Total	71	3,55

Tabela 6: Desempenho da árvore de decisão em relação à variável “Entra ou não na Loja” no dia 9 de Setembro

Assim a taxa de erro deste modelo é de 3,55%: classifica incorrectamente 10 casos com a classe “entra” quando na verdade não entram e classifica incorrectamente 61 casos com a classe “não entra” quando na verdade entram.

Em relação à distribuição dos dados pelas duas classes, das 1998 observações, cerca de 56% pertence à classe “não entra” (1121 casos) e cerca de 44% à classe “entra” (877 casos). Assim a taxa de erro é maior para a classe “Entra” do que para a classe “Não Entra”, cerca de 7% face a cerca de 0,9%, respectivamente.

Segundo o modelo obtido:

- A primeira variável usada para classificar se um indivíduo entra ou não na loja é “Tempo de Observação da Montra”: se o indivíduo não observar a montra (“Tempo de Observação da Montra” ≤ 0) ele irá entrar na loja (ponto (1) da Figura n.º 2). São assim classificados 796 casos, ou seja, aproximadamente 40% dos casos;
- Se o Indivíduo observar a montra (“Tempo de Observação da Montra” > 0), ele irá entrar ou não na loja dependendo do movimento em frente da montra:
 - Se o indivíduo não passar em frente da montra (“Movimento Montra” = 0), não entra na loja (ponto (2) da Figura n.º 2);
 - Se permanecer quieto em frente da montra (“Movimento Montra” = 1), entrará na loja consoante pára ou não em frente da montra: se não parar (“NãoPáraVsPára” = 0) ele entra (ponto (3) da Figura n.º 2), se parar (“NãoPáraVsPára” = 1), entra ou não consoante o ataque montra. Se não houver ataque á montra ou se este for natural (“Ataque Montra” = 0 e “Ataque Montra” = 1, respectivamente), o indivíduo não entra (ponto (4) da Figura n.º 2), caso contrário a decisão dependerá do comportamento em frente da montra.
 - ✓ Assim se não se registar qualquer comportamento em frente à montra, ou seja se o indivíduo não observar a montra (“Comportamento Montra” = 0), ele entra (ponto (5) da Figura n.º 2);
 - ✓ Se fixar-se à direita (“Comportamento Montra” = 1) e o sexo do indivíduo/grupo for masculino (“Sexo Grupo” = M), entra (ponto (6) da Figura n.º 2), se ambos os sexos estiverem presente no grupo (“Sexo

Grupo” = Misto), este entrará caso a média da idade for menor ou igual a 2 (ponto (7) da Figura n.º 2), caso contrário não entra (ponto (8) da Figura n.º 2). Se o sexo do indivíduo/grupo for feminino (“Sexo Grupo” = F), entra, caso venha da direita (“Comportamento Corredor” = 1), isto é, caso não necessite de passar em frente da montra para entrar (ponto (9) da Figura n.º 2), se vier da esquerda (“Comportamento Corredor” = 2) entrará caso o tempo de observação da montra seja inferior ou igual a cerca de 12 segundos (ponto (10) da Figura n.º 2), caso contrário não entra (ponto (11) da Figura n.º 2);

✓ Se fixar-se ao centro (“Comportamento Montra” = 2) e o sexo do indivíduo/grupo for masculino (“Sexo Grupo” = M), não entra (ponto (12) da Figura n.º 2), se o sexo do indivíduo/grupo for feminino ou se ambos os sexos estiverem presente no grupo (“Sexo Grupo” = F e “Sexo Grupo” = Misto, respectivamente), entrará (ponto (13) da Figura n.º 2);

✓ Se fixar-se à esquerda (“Comportamento Montra” = 3) o indivíduo/grupo não entra (ponto (14) da Figura n.º 2);

✓ Se fixar-se em mais do que um ponto da montra (“Comportamento Montra” = 4), o indivíduo/grupo entra (ponto (15) da Figura n.º 2);

- Se o indivíduo/grupo se movimenta em frente da montra (“Movimento Montra” = 2), entrará na loja consoante o tempo de observação da montra: se for menor ou igual a cerca de 9 segundos não entra (ponto (16) da Figura n.º 2), caso contrário, se não parar em frente da montra (“NãoPáraVsPára” = 0), não entra (ponto (17) da Figura n.º 2), se parar (“NãoPáraVsPára” = 1) e vier da direita (“Deslocação Corredor = 1), entra (ponto (18) da Figura n.º 2), se vier da esquerda (“Deslocação Corredor = 2), só entra se a média de idades for inferior ou igual a cerca de, 2,6 (ponto (19) da Figura n.º 2), caso contrário não entra (ponto (20) da Figura n.º 2).

Resumindo, a primeira variável tida em consideração no método de classificação é o tempo de observação da montra, e a variável considerada posteriormente é a variável “Movimento em Frente da Montra”, pelo que só se entra se registar-se algum

movimento ou se permanecer quieto a observar a montra. Este modelo tem maior probabilidade de erro ao classificar quem entra na loja, pois a taxa de erro obtida para esta classe é maior.

5.1.2. MODELO DE CLASSIFICAÇÃO OBTIDO PARA A VARIÁVEL “ENTRA OU NÃO NA LOJA” NO DIA 16 DE SETEMBRO

Numa segunda fase, o modelo de classificação foi construído usando 2051 observações que correspondem às recolhidas no dia 16 de Setembro. Obteve-se o seguinte modelo:

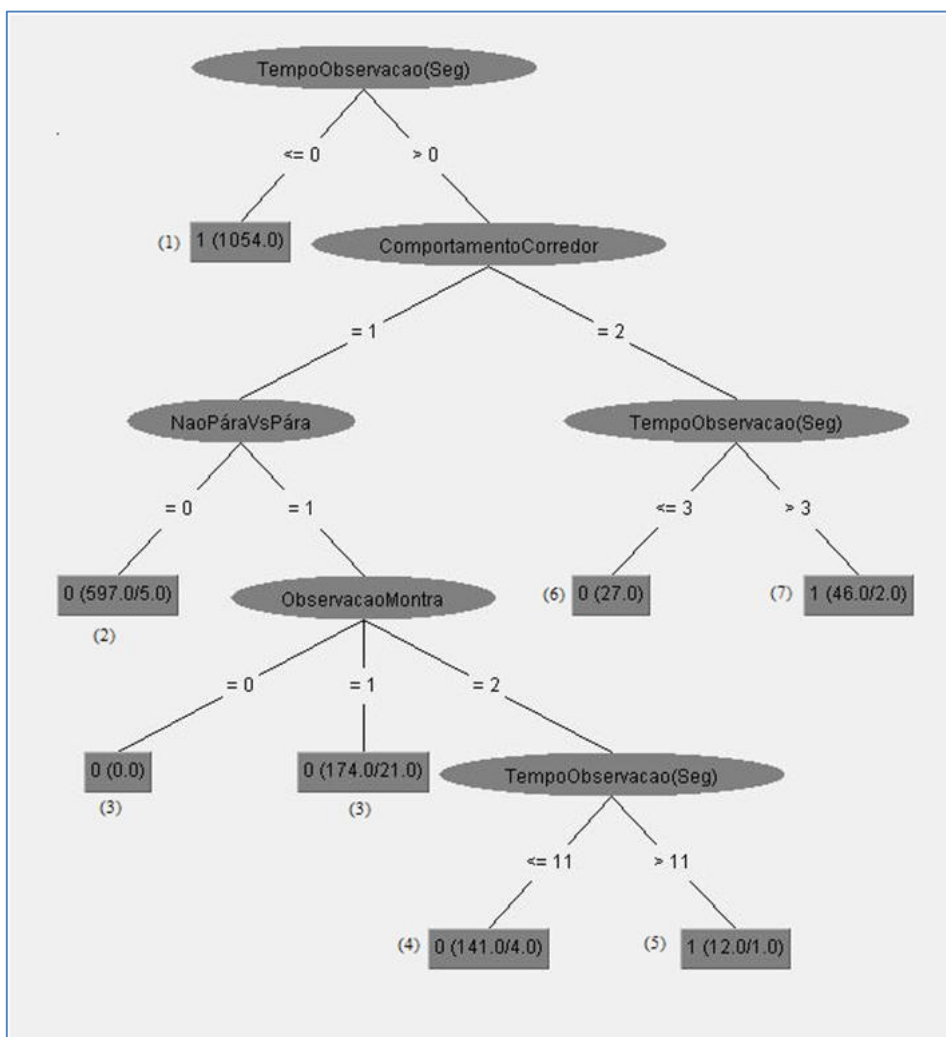


Figura n.º 3: Árvore de Decisão “Entra ou não na loja” no dia 16 de Setembro

Este modelo classifica correctamente 2016 casos como podemos constatar no output gerado pelo modelo J48 do Weka do Anexo n.º 3, resumido na Tabela 7.

		N.º	%
Casos Classificados Correctamente		2016	98,29
Casos Classificados Incorrectamente	Como “Entram na Loja”	3	0,15
	Como “Não Entram na Loja”	32	1,56
	Total	35	1,71

Tabela 7: Desempenho da árvore de decisão em relação à variável “Entra ou não na Loja” no dia 16 de Setembro

Assim a taxa de erro deste modelo é de 1,71%: classifica incorrectamente 3 casos com a classe “entra” quando na verdade não entram e classifica incorrectamente 32 casos com a classe “não entra” quando na verdade entram.

Em relação à distribuição dos dados pelas duas classes, das 2051 observações, cerca de 44% pertence à classe “não entra” (912 casos) e cerca de 56% à classe “entra” (1139 casos). Assim a taxa de erro é mais uma vez é maior para a classe “Entra” do que para a classe “Não Entra”, cerca de 2,8% face a cerca de 0,33%, respectivamente.

Novamente a primeira variável usada para classificar se um indivíduo/grupo entra ou não na loja é “Tempo de Observação da Montra”: se o indivíduo/grupo não despende tempo na observação da montra (“Tempo de Observação da Montra” ≤ 0), ele irá entrar na loja (ponto (1) da Figura n.º 3). São assim classificados 1054 casos, isto é, cerca de 51% dos casos em estudo. Se o Indivíduo despende tempo a observar a montra (“Tempo de Observação da Montra” > 0), ele irá entrar ou não na loja dependendo do comportamento no corredor:

- Se vem da direita (“Comportamento Corredor” = 1) e não pára para ver a montra (“Não Pára Vs Pára” = 0), não entra (ponto (2) da Figura n.º 3), se pára para ver a montra (“Não Pára Vs Pára” = 1), mas não tem qualquer comportamento de observação da montra ou foca um dado aspecto da montra (“Observação Montra” = 0 e “Observação Montra” = 1, respectivamente), também não entra (ponto (3) da Figura n.º 3), se ao observar a montra a varre (“Observação Montra” = 2), então

entra, caso permaneça mais de cerca de 11 segundos a observa-la (ponto (4) da Figura n.º 3), caso contrário não entra (ponto (5) da Figura n.º 3);

- Se vem da esquerda, isto é, passa necessariamente em frente da montra (“Comportamento Corredor” = 2), o indivíduo entra, se permanecer mais de cerca de 3 segundos a observa-la (ponto (6) da Figura n.º 3), caso contrário não entra (ponto (7) da Figura n.º 3).

Resumindo a primeira variável tida em consideração neste método de classificação é também o tempo de observação da montra, no entanto, a variável considerada posteriormente já é diferente da registada no dia 9: neste dia é considerada a variável “Movimento em Frente da Montra”, e no dia 16 a variável “Comportamento no Corredor”, isto é, se os consumidores vêm da direita, caso em que para entrar não necessita observar a montra; ou da esquerda, caso em que passa necessariamente pela montra e só depois pela porta da loja. Mais uma vez o modelo obtido tem maior probabilidade de erro ao classificar quem entra na loja, pois a taxa de erro obtida para esta classe é maior.

5.2. QUEM COMPRA

Para analisar quem compra apenas foram consideradas as observações de quem entra na loja, pois só nestes casos poderá haver compra. Deste modo, apenas foram consideradas 877 observações no dia 9 de Setembro e 1139 no dia 16 de Setembro.

Pretende-se saber em que medida o comportamento face à montra permite classificar se um indivíduo compra ou não.

Nesta análise não foram consideradas as seguintes variáveis:

- Número de sacos de Compras;
- Tipo de Compra;
- Quem Compra.

A exclusão destas variáveis resulta do facto de todas serem consequência de o indivíduo comprar ou não. Por simplificação, uma vez que se inclui o Tempo de Observação da Montra e o Tempo na Loja, não foram consideradas as variáveis Hora de Início da Observação, Hora de Fim da Observação; Hora Entrada da Loja, e Hora Saída da Loja.

Para analisar quem compra vai-se então aplicar o método J48 para se obter Árvores de Decisão em relação à variável “Compra ou Não”. Tal como anteriormente, isto será feito numa primeira fase para as observações recolhidas no dia 9 de Setembro e numa segunda fase para as observações recolhidas no dia 16 de Setembro, e em ambos os casos considerando apenas quem entra na loja.

5.2.1. MODELO DE CLASSIFICAÇÃO OBTIDO PARA A VARIÁVEL “COMPRA OU NÃO” NO DIA 9 DE SETEMBRO

Numa primeira fase obteve-se o modelo de árvore de classificação para o dia 9 de Setembro, considerando as 877 observações, o qual é apresentado na Figura n.º 4.

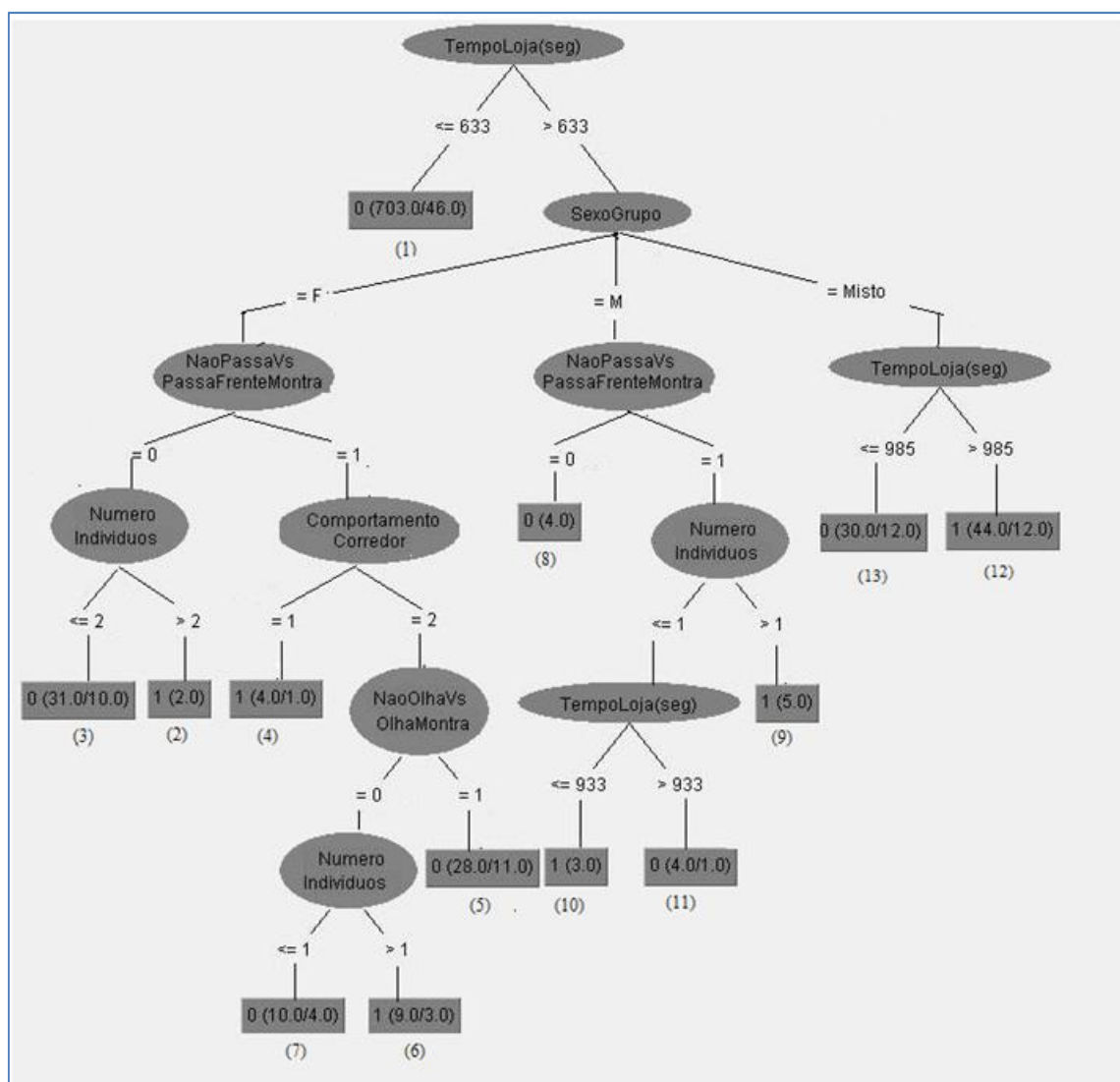


Figura n.º 4: Árvore de Decisão “Compra ou não” no dia 9 de Setembro

Este modelo classifica correctamente 734 casos como podemos constatar no output gerado pelo modelo J48 do Weka do Anexo n.º 4, resumido na Tabela 8.

		N.º	%
Casos Classificados Correctamente		734	83,69
Casos Classificados Incorrectamente	Como “Compra”	49	5,59
	Como “Não Compra”	94	10,71
	Total	143	16,30

Tabela 8: Desempenho da árvore de decisão em relação à variável “Compra ou não” no dia 9 de Setembro

Assim a taxa de erro deste modelo é de 16,3%: classifica incorrectamente 49 casos com a classe “compra” quando na verdade não compram e classifica incorrectamente 94 casos com a classe “não compra” quando na verdade compra. Em relação à totalidade dos casos que pertencem à classe “compra” (135 casos) o modelo classifica mais casos de forma errada do que correcta (94 e 41 casos respectivamente).

Em relação à distribuição dos dados pelas duas classes, das 877 observações, cerca de 85% pertence à classe “não compra” (742 casos) e cerca de 15% à classe “compra” (135 casos). Assim a taxa de erro é maior para a classe “compra” do que para a classe “Não Compra”, cerca de 69,63% face a cerca de 6,6%, respectivamente.

A primeira variável considerada para classificar se o indivíduo compra ou não é o tempo de permanência na loja: se o tempo de permanência na loja for menor ou igual a cerca de 11 minutos (“Tempo Loja (seg)” $\leq$ 633), não compra (ponto (1) da Figura n.º 4), caso contrário comprará ou não dependendo do Sexo do indivíduo/grupo.

Este comportamento seria de esperar porque subjacente ao acto de comprar está o de experimentar o bem que se pretende adquirir, o que torna o tempo de permanência na loja maior.

Relativamente ao sexo:

- Se o indivíduo/grupo for do sexo feminino (“Sexo Grupo” = F), e o indivíduo não passa em frente da montra (“Não passa Vs Passa Frente Montra” = 0), compra se o grupo for constituído por mais do que 2 pessoas (“Número Indivíduos” $>$ 2) (ponto (2) da Figura n.º 4), caso contrário não compra (ponto (3) da Figura n.º 4). Se o indivíduo passar em frente da montra (“Não passa Vs Passa Frente Montra” = 1) e o indivíduo/grupo vem da direita, pelo que para entrar na montra não necessita passar em frente dela (“Comportamento Corredor” = 1), ele compra (ponto (4) da Figura n.º 4), se vem da esquerda (“Comportamento Corredor” = 2), e olha para a montra (“Não Olha Vs Olha Montra” = 1), não compra (ponto (5) da Figura n.º 4), se não olha para a montra (“Não Olha Vs Olha Montra” = 0), compra se a observação for composta por mais do que um indivíduo (“Número Indivíduos” $>$ 1) (ponto (6) da Figura n.º 3), caso contrário não compra (ponto (7) da Figura n.º 4);

- Se o indivíduo/grupo for do sexo masculino (“Sexo Grupo” = M), e o indivíduo não passa em frente da montra (“Não passa Vs Passa Frente Montra” = 0), não há compra (ponto (8) da Figura n.º 4). Se o indivíduo passa em frente da montra (“Não passa Vs Passa Frente Montra” = 1) e a observação for composta por mais do que um indivíduo (“Número Indivíduos” > 1), há compra (ponto (9) da Figura n.º 3), caso contrário isto é, se for um único indivíduo do sexo masculino às compras, existirá compra se o tempo na loja for inferior a cerca de 16 minutos (ponto (10) da Figura n.º 4), caso contrário não (ponto (11) da Figura n.º 4).
- Se o grupo for composto pelos dois sexos (“Sexo Grupo” = Misto), haverá compra se permanecerem na loja mais de cerca de 16 minutos (ponto (12) da Figura n.º 4), caso contrário não (ponto (13) da Figura n.º 3).

Através dos resultados verifica-se que, além do tempo de permanência na loja, parece influenciar a compra o sexo dos indivíduos do grupo que se encontra às compras. O modelo obtido tem maior probabilidade de erro ao classificar quem compra, pois a taxa de erro obtida para esta classe é maior.

5.2.2. MODELO DE CLASSIFICAÇÃO OBTIDO PARA A VARIÁVEL “COMPRA OU NÃO” NO DIA 9 DE SETEMBRO

Numa segunda fase para o dia 16 de Setembro, considerando 1139 casos, obteve-se o modelo da Figura n.º 5.

Segundo este modelo um indivíduo/grupo se permanecer na loja menos de cerca de 8 minutos (“Tempo Loja (seg)” ≤ 494), não efectuará compra (ponto (1) da Figura n.º 5). No entanto, se demorar dentro da loja mais de cerca de 8 minutos, a decisão dependerá do número de pessoas que compõem a observação:

- Se a observação for composta por um ou dois indivíduos (“Número Indivíduos” ≤ 2)

- Se o indivíduo/grupo pára em frente da montra (“Não Pára Vs Pára” = 1), e se deslocar no lado oposto à montra (“Deslocação Corredor” = 2), realizará a compra (ponto (2) da Figura n.º 5). Se deslocar-se no lado da montra, compra se observar a montra mais de cerca de 7 segundos (ponto (3) da Figura n.º 5), caso contrário não compra (ponto (4) da Figura n.º 5).

- Se o indivíduo/grupo não pára em frente da montra (“Não Pára Vs Pára” = 0):

- ✓ Se deslocar-se no lado da montra (“Deslocação Corredor” = 1), compra, se permanecer na loja mais de cerca de 20 minutos (ponto (5) da Figura n.º 5). Caso contrário a decisão depende do movimento em frente da montra: não compra (ponto (6) da Figura n.º 5) se permanece quieto em frente da montra ou se movimentar-se ao longo dela (“Movimento Montra” = 1 e “Movimento Montra” = 2, respectivamente), mas se não observar a montra, não se registando qualquer movimento em frente dela (“Movimento Montra” = 0), compra se de a média das idades que compõem o grupo for maior de 2,25 (ponto (7) da Figura n.º 5), caso contrário não compra (ponto (8) da Figura n.º 5);

- ✓ Se deslocar-se no lado oposto à montra (“Deslocação Corredor” = 2), se vier da direita (“Comportamento no corredor” = 1) e se estiver dentro da loja mais de cerca de 10 minutos, compra (ponto (9) da Figura n.º 5), caso contrário não compra (ponto (10) da Figura n.º 5). Se vier pela esquerda (“Comportamento no corredor” = 2), compra dependendo do sexo de indivíduo de cada observação: Se for composta exclusivamente pelo sexo feminino, compra se estiver dentro da loja menos de cerca de 10 minutos (ponto (11) da Figura n.º 5), caso

contrário não compra (ponto (12) da Figura n.º 5); se for composta exclusivamente pelo sexo masculino compra se permanecerem na loja entre cerca de 12 e 16 minutos (ponto (13) da Figura n.º 5), caso contrário não compra (ponto (14) da Figura n.º 5); se a observação for composta pelos dois sexos, compra se permanece na loja mais de cerca de 13 minutos (ponto (15) da Figura n.º 5) caso contrário não compra (ponto (16) da Figura n.º 5);

▪ Se a observação for composta por mais de dois indivíduos (“Número Indivíduos” > 2):

- Se forem exclusivamente do sexo feminino compram (ponto (17) da Figura n.º 5);
- Se forem exclusivamente do sexo masculino não compram (ponto (18) da Figura n.º 5);
- Se a observação for composta pelos dois sexos, caso parem em frente da montra (“Não Pára Vs Pára” = 1), não compram (ponto (19) da Figura n.º 5). Se não pararem em frente da montra (“Não Pára Vs Pára” = 0), compram se estiverem na loja mais de cerca de 23 minutos (ponto (20) da Figura n.º 5). Caso contrário a compra depende do número de indivíduos da observação: se o grupo for composto por mais de três indivíduos, compram se estiverem na loja entre cerca de 9 e 11 minutos (ponto (21) da Figura n.º 5), caso contrário não compram (ponto (22) da Figura n.º 5); se o grupo for composto por três indivíduos haverá compra se não tiverem olhado para a montra e a média de idades for inferior a 2,75 (ponto (23) da Figura n.º 5), caso contrário não compram (ponto (24) da Figura n.º 5).

Através dos resultados obtidos, além do tempo de permanência na loja, parece influenciar a compra o número de indivíduos do grupo que se encontra às compras. O modelo obtido tem maior probabilidade de erro ao classificar quem compra, pois a taxa de erro obtida para esta classe é maior.

Obtiveram-se nesta secção os quatro modelos de classificação dos consumidores: Entra ou Não Entra na Loja no dia 9 de Setembro; Entra ou Não Entra na Loja no dia 16 de

Setembro; Compra ou Não no dia 9 de Setembro; Compra ou Não no dia 16 de Setembro. De seguida será apresentada uma discussão dos resultados aqui apresentados.

6. DISCUSSÃO DOS RESULTADOS E APLICAÇÃO NO NEGÓCIO

Os resultados obtidos permitem classificar se um indivíduo/grupo entra ou não na loja, ou se realiza ou não uma compra em resultado do seu comportamento face á montra. Esta análise tem uma grande importância na medida em que a descoberta do que leva o consumidor a entrar e a comprar permitirá tomar medidas no sentido de aumentar o poder de atracção da loja sobre consumidores, bem como o volume de vendas.

Para classificar quem entra na loja a primeira variável considerada é o “Tempo de Observação da Montra”, nos dois dias de observação. De facto, nos períodos pré-saldos e pós-saldos, se o indivíduo/grupo não despende tempo a observar a montra, ele entra na loja em 39,84% e 51,39% dos casos, respectivamente. Isto poderá indicar que existe uma percentagem considerável de pessoas que entram na loja independentemente do poder de atracção da montra: podem conhecer a marca, ser consumidores habituais, e já ter a intenção de entrar. Estes casos poderão assim resultar de o consumidor ter planeado a ida à loja

Dado o elevado número de casos registados nesta situação seria interessante analisar o que motivou, em cada observação, a entrada na loja: terá sido alguma campanha de marketing ou de publicidade? Será o poder de atracção e o reconhecimento da marca? Esta informação seria útil para se poder actuar e atrair cada vez mais potenciais consumidores.

Um conjunto considerável de casos classificados como “Não Entra na loja” no dia 9 de Setembro, regista-se para aqueles que despendem tempo a observar a montra, se movimentam ao longo dela, mas o tempo de observação da montra não ultrapassa cerca de 9 segundos. São assim classificados 46,10% dos casos relativos ao dia 9 de Setembro. Este resultado poderá sugerir que a montra deveria ter sido capaz de atrair a atenção dos potenciais consumidores por mais de cerca de 9 segundos: neste caso,

segundo o modelo obtido, o indivíduo/grupo entraria na loja se parasse a observar a montra.

No dia 16 de Setembro, 29,11% dos casos são classificados como “Não Entram” na loja, se vierem da direita e apesar de despenderem algum tempo a observar a montra, não param para a observar. Estes indivíduos, para entrarem na loja, teriam de voltar atrás na sua trajectória depois de observar a montra. Assim, apenas a observam à medida que continuam o seu caminho. Também neste dia a montra deveria ter sido capaz de ter um maior efeito de atracção sobre os consumidores de modo a “convencê-los” a voltar atrás e a entrar.

Pode-se, para os dois dias, resumir em que casos os indivíduos entram na loja.

No dia 9 de Setembro entram na loja:

- Quem não despende qualquer tempo a observar a montra;
- Quem despende tempo a observar a montra; permanece quieto em frente dela, pára para a observar; registando-se um ataque de aproximação à montra e:
 - Fixa-se em mais do que um ponto da montra;
 - Fixa-se ao centro e caso o grupo seja composto pelo sexo feminino ou misto
 - Fixa-se à direita e ou é do sexo masculino ou é um grupo composto pelos dois sexos e a média das idades é menor que 2 ou é do sexo feminino e vem da direita ou é do sexo feminino vem da esquerda e observa a montra menos de cerca de 12 segundos.
- Quem despende tempo a observar a montra; faz pequenas deslocações em frente da mesma com o objectivo de a observar, pelo que esse tempo de observação terá de ser mais de 9 segundos, pára em frente dela, e:
 - Vem do lado direito (posição do indivíduo transeunte em relação à montra), isto passa, primeiro em frente da porta da loja do que da montra, pelo que antes de entrar, vai intencionalmente observar a montra.
 - Vem do lado esquerdo da montra, isto é para entrar na loja passa obrigatoriamente em frente da montra, mas a média das idades das pessoas

que compõem o grupo, é inferior a 2,6, isto é, é maioritariamente constituído por pessoas adultas ou jovens.

Existe um grupo de indivíduos para o qual a montra é importante, uma vez que a decisão de entrar na loja parece ser motivada pela mesma, dada o comportamento que assume em frente dela; mas existe um outro grupo de indivíduos que entra independentemente da montra, nem sequer parando para a observar, sendo este o maior número de casos.

No dia 16 de Setembro entram na loja:

- Quem não despende qualquer tempo a observar a montra;
- Quem despende mais de cerca de 3 segundos a observar a montra e vem da esquerda;
- Quem despende mais de cerca de 11 segundos a observar a montra, vem da direita, pára para a observar e varre a montra. Estes indivíduos se observam a montra, têm de voltar atrás para entrar, pelo que demonstram uma maior influenciam da montra na decisão de entra.

Mais uma vez temos aqueles que manifestam interesse pela montra e entram e aqueles que não param sequer para a observar, sendo mais uma vez este o maior número de casos.

Analisando quem compra, a primeira variável usada para classificar um caso é o tempo na loja, nos dois dias observados, classificando que o indivíduo/grupo não compra se esse tempo for inferior a cerca de 11 e 8 minutos, respectivamente. De facto é de esperar que haja uma relação directa entre a compra e o tempo dispendido na loja, uma vez que é de esperar que quem compra experimente, logo demore mais tempo na loja.

A segunda variável usada para classificar os indivíduos que permanecem na loja mais tempo que o anteriormente referido é o Sexo do indivíduo/grupo no dia 9 de Setembro e o Número de indivíduos do grupo no dia 16 de Setembro. Assim, no dia 9 de Setembro o maior número de casos classificado como “compra” regista-se para os grupos constituídos pelos dois sexos, que permanecem na loja mais de cerca de 16 minutos. No

dia 16 de Setembro, o maior número de casos classificados como “compra” registam-se para as observações compostas por um ou dois indivíduos.

Vamos então analisar quem compra em cada um dos dias.

No dia 9 de Setembro compra:

- Os grupos de indivíduos compostos pelos dois sexos, que permanecem na loja mais de cerca de 16 minutos;
- Os grupos de indivíduos compostos por mais do que um indivíduo, mas apenas do sexo masculino, que passam em frente da montra e permanecem na loja mais de cerca de 11 minutos;
- Os grupos de indivíduos compostos apenas por um indivíduo do sexo masculino, que passam em frente da montra e permanece na loja entre cerca de 11 e 16 minutos;
- Os grupos de indivíduos compostos por mais do que dois indivíduos, mas apenas do sexo feminino, que não passam em frente da montra, e permanecem na loja mais de cerca de 11 minutos;
- Os grupos de indivíduos compostos por um ou mais indivíduos, apenas do sexo feminino, que passam em frente da montra, vêm da direita e permanecem na loja mais de cerca de 11 minutos;
- Os grupos de indivíduos compostos por mais do que um indivíduo, apenas do sexo feminino, que passam em frente da montra, vêm da esquerda, não olham para a montra e permanecem na loja mais de cerca de 11 minutos.

No dia 9 de Setembro é assim relevante na decisão de comprar ou não, o tempo que o indivíduo permanece na loja, o número de indivíduos que compõem o grupo e o sexo dos mesmos. A avaliar pelos resultados obtidos a compra verifica-se mais quando os indivíduos estão acompanhados.

Para avaliar o comportamento de quem compra no dia 9 de Setembro, em relação à montra, apenas são utilizadas as variáveis que permitem constatar se o grupo passa ou não em frente da montra, se vem da direita ou da esquerda do corredor e se olha ou não para a montra. Através da análise destas variáveis verifica-se que no dia 9 de Setembro

a montra parece não ter um efeito significativo da decisão de comprar, provavelmente porque é época de saldos, e logo as pessoas deslocam-se à loja com o objectivo de entrarem para ver os artigos em saldos, à procura de bons preços, independentemente do exposto na montra.

No dia 16 de Setembro compram:

- Os grupos compostos por mais de 2 indivíduos do sexo feminino, que permanecem na loja mais de cerca de 8 minutos;
- Os grupos compostos por mais de 2 indivíduos dos dois sexos, que não param em frente da montra e permanecem na loja mais de cerca de 43 minutos;
- Os grupos compostos por mais de 3 indivíduos dos dois sexos, que não param em frente da montra e permanecem na loja entre cerca de 7 e 11 minutos;
- Os grupos compostos por 3 indivíduos dos dois sexos, que não param em frente da montra, não olham para a montra, permanecem na loja menos de cerca de 43 minutos e a média de idade é inferior a 2,75;
- As observações compostas por um ou dois indivíduos, que param em frente da montra, deslocam-se no lado oposto à montra e permanecem na loja mais de cerca de 8 minutos;
- As observações compostas por um ou dois indivíduos, que param em frente da montra, deslocam-se no lado da montra, observam-na mais de 7 segundos e permanecem na loja mais de cerca de 8 minutos;
- As observações compostas por um ou dois indivíduos, que não param em frente da montra, deslocam-se no lado da montra e permanecem na loja mais de cerca de 20 minutos;
- As observações compostas por um ou dois indivíduos, que não param em frente da montra, deslocam-se no lado da montra, não têm qualquer movimento em frente da montra, permanecem na loja entre cerca de 8 e 20 minutos, e a média de idades é maior que 2,25;
- As observações compostas por um ou dois indivíduos, que não param em frente da montra, deslocam-se no lado oposto da montra, vêm da direita e permanecem na loja entre cerca de 8 e 10 minutos;

- As observações compostas por um ou dois indivíduos do sexo feminino, que não param em frente da montra, deslocam-se no lado oposto da montra, vêm da esquerda, e permanecem na loja entre cerca de 8 e 10 minutos;
- As observações compostas por um ou dois indivíduos do sexo masculino, que não param em frente da montra, deslocam-se no lado oposto da montra, vêm da esquerda, e permanecem na loja entre cerca de 12 e 16 minutos;
- As observações compostas por dois indivíduos, um de cada sexo, que não param em frente da montra, deslocam-se no lado oposto da montra, vêm da esquerda, olham para a montra e permanecem na loja mais de cerca de 14 minutos.

Face aos resultados obtidos poderemos agora analisar em que medida a decisão de entrar e de comprar é ou não influenciada por variáveis diferentes consoante estamos ou não na época de saldos.

Para a comparação dos resultados obtidos na época de saldos e pós-saldos, interessa referir que a montra da Salsa, nos dois momentos de recolha de informação estava construída da mesma forma: um modelo feminino do lado direito da montra (posição do transeunte face à montra), um modelo feminino e um modelo masculino do lado esquerdo e, ao centro, algumas peças de roupa com especial destaque (no dia 9 estavam pousadas sobre uns móveis suspensos do tecto, no dia 16 pousadas sobre umas mesas, logo nos dois dias à mesma altura dos olhos do transeunte).

Esta informação é relevante na medida em que o impacto da montra sob o transeunte poderia derivar da forma como a mesma está construída, em vez do efeito que se pretende estudar que é o efeito dos saldos. De facto existem estudos que demonstram que os artigos colocados à altura dos nossos olhos são mais facilmente percebidos pelos indivíduos, por outro lado também as cores são importantes para atrair a atenção [3].

O que se pretende analisar é o eventual efeito da época de saldos na decisão de entrar na loja e de comprar o produto. Este método de promoção de vendas é um instrumento que é cada vez mais utilizado. Tal resulta do facto de hoje as empresas a retalho terem um grande poder face às empresas grossistas, que lhe confere mais autonomia face ao passado; de a lealdade à marca ser uma característica cada vez mais rara entre os

consumidores; de os consumidores responderem positivamente a este incentivo de compra; do mercado estar cada vez mais segmentado; da necessidade de obter receitas no curto prazo e de ser um método mais económico face, por exemplo, a publicidade [13].

Comparando os resultados apresentados para a decisão de entrar ou não da loja no dia de saldos e pós saldos, verifica-se que o comportamento face à montra assume maior relevância no dia 16 de Setembro para classificar que um indivíduo entra na loja, uma vez que nesse dia existe uma maior proporção de casos classificados como entra na loja em função desse comportamento: 58 casos são classificados como entram na loja no dia 16 de Setembro em função desse comportamento face a 33 casos no dia 9 de Setembro. Isto poderá resultar do facto de a época de saldos por si só poder ser um incentivo para a entrada na loja, pois as pessoas podem ter o objectivo de entrar para ver os artigos em saldos, à procura de bons preços, independentemente do exposto na montra; enquanto no dia 16, onde já é apresentada a nova colecção, a montra assume maior importância no sentido de dar a conhecer as novidades, e por tal é objecto de maior atenção.

Assim, julgo poder-se subentender que a época de saldos é um incentivo à entrada na loja, e que a montra é um instrumento de atracção e captação de potenciais clientes, mais relevante quando está a ser apresentada a nova colecção, isto é quando é um veículo de informação para os consumidores. Na época de saldos não se está à espera que a montra apresente novidades, pois os artigos já são conhecidos e o incentivo à compra está associado ao preço.

Em relação à decisão de comprar, no dia 9 de Setembro é relevante o tempo que o indivíduo permanece na loja, o número de indivíduos que compõem o grupo e o sexo dos mesmos. Para avaliar o comportamento de quem compra em relação à montra no dia 9 de Setembro, apenas são utilizadas as variáveis que permitem constatar se o grupo passa ou não em frente da montra, se vem da direita ou da esquerda do corredor e se olha ou não para a montra.

No dia 16 de Setembro, para avaliar esse comportamento, para além das variáveis também usadas no modelo obtido para a época de saldos, são também usadas as variáveis para ou não em frente da montra e movimento em frente da montra, pelo que

subentende-se que neste caso a montra tem um maior impacto na decisão de compra dos indivíduos.

Assim, como já foi referido, na decisão de comprar, as variáveis sociais (sexo, número indivíduos grupo, idade) demonstram mais relevância do que as situacionais (impacto da montra). Ainda assim o comportamento face à montra parece demonstrar uma maior influência na decisão de comprar no dia 16 de Setembro, pois são consideradas mais variáveis que permitem descrever esse comportamento e são classificados mais casos em função dessas variáveis. A montra revela-se assim um instrumento mais eficaz de captação de clientes na época da nova colecção, pois os consumidores demonstraram maior interesse pelos artigos da nova colecção Outono-Inverno.

Nesta secção pode-se assim avaliar do impacto da montra enquanto agente motivador da entrada na loja e da compra. Concluiu-se que a maioria dos indivíduos é classificada como entra na loja sem observar a montra e que a maioria dos indivíduos é classificada como compradora caso permaneça na loja por um certo período de tempo, consoante o sexo e o número de indivíduos que fazem parte da observação do dia 9 ou 16 de Setembro, respectivamente. Finalmente concluiu-se que os consumidores demonstram maior interesse pela montra, quando esta apresenta a nossa colecção.

7. CONCLUSÃO

A presente dissertação pretendeu analisar em que medida o comportamento de um consumidor face a uma montra nos permite aferir se esse consumidor entra ou não na loja e se compra ou não, bem como analisar em que medida esse comportamento é diferente consoante estamos ou não na época de saldos. Para tal utilizou um conjunto de dados, recolhidos nos dias 9 e 16 de Setembro de 2006, respectivamente na época de saldos e na época de nova colecção, na loja Salsa situada no Norte Shopping, através de duas câmaras de filmar (uma câmara de vídeo exterior permitiu captar os consumidores que se deslocavam no exterior da loja; uma câmara de vídeo interior, permitiu analisar se os consumidores realizavam ou não compras).

Pretendeu-se analisar em que medida a montra, entre os diferentes instrumentos ao alcance do retalhista para atrair e incentivar o consumidor, pode ou não ser relevante para captar a sua atenção, incentivando-o a entrar e a comprar, uma vez que é o seu primeiro contacto com a loja: poderá conhecer a marca ou os artigos que estão à venda, mas é a montra que lhe dá uma primeira impressão da loja em si.

Iremos então apresentar os principais resultados e as principais conclusões obtidas com este estudo, que permitem responder às questões colocadas e que são a essência desta abordagem: Qual a influência da montra na decisão de entrar na loja? Quais as variáveis mais importantes na decisão de compra? A decisão de entrar e de comprar é ou não influenciada por variáveis diferentes consoante estamos ou não na época de saldos?

7.1. RESULTADOS

O método de classificação usado neste estudo foi o método das Árvores de Decisão, as quais foram obtidas através do programa Weka. Foram assim construídos quatro modelos de classificação: Entra ou Não Entra na Loja no dia 9 de Setembro; Entra ou

Não Entra na Loja no dia 16 de Setembro; Compra ou não no dia 9 de Setembro e Compra ou não no dia 16 de Setembro. Vamos então apresentar as conclusões de cada um dos modelos obtidos.

7.1.1. QUAL A INFLUÊNCIA DA MONTRA NA DECISÃO DE ENTRAR OU NÃO NA LOJA?

Para classificar quem entra na loja a primeira variável considerada é o tempo de observação da montra, nos dois dias de observação. De facto, quer no período pré-saldos, quer no período pós-saldos, se o indivíduo/grupo não despende tempo a observar a montra, ele entra na loja em 39,84% e 51,39% dos casos, respectivamente. Isto poderá indicar que existe uma percentagem considerável de pessoas que entram na loja independentemente do poder de atracção da montra: podem conhecer a marca, ser consumidores habituais, já ter a intenção de entrar. Estes casos poderão assim derivar de o consumidor ter planeado a ida à loja.

Um conjunto considerável de casos classificados como “não entra na loja” no dia 9 de Setembro, regista-se para aqueles que despendem tempo a observar a montra, se movimentam ao longo dela, mas o tempo de observação da montra não ultrapassa os cerca de 9 segundos. São assim classificados 46,10% dos casos. Este resultado poderá sugerir que a montra deveria ter sido capaz de atrair a atenção dos potenciais consumidores por mais de cerca de 9 segundos.

No dia 16 de Setembro 29,11% dos casos são classificados que “não entram” na loja se vierem da direita e apesar de despendem algum tempo a observar a montra, não param para a observar. Estes indivíduos para entrarem na loja teriam de voltar atrás na sua trajectória depois de observar a montra. Assim, apenas a observam à medida que continuam o seu caminho. Também neste dia a montra deveria ter sido capaz de ter um maior efeito de atracção sobre os consumidores de modo a “convence-los” a voltar atrás e a entrar.

Assim, entra na loja no dia 9 de Setembro quem não despende qualquer tempo a observar a montra; quem despende tempo a observar a montra, permanece quieto em frente dela mas não pára para a observar (uma vez que se parar a decisão de entrar já dependerá do ataque à montra e do comportamento em frente à montra); quem despende mais de 9 segundos a observar a montra, faz pequenas deslocações em frente da mesma, pára em frente dela, e vem do lado direito (isto é, passa primeiro em frente da porta da loja do que da montra, pelo que antes de entrar, vai intencionalmente observar a montra), se vier do lado esquerdo da montra, isto é, para entrar na loja passa obrigatoriamente em frente da montra, a decisão de entrar dependerá da idade das pessoas que observam a montra.

No dia 16 de Setembro entram na loja quem não despende qualquer tempo a observar a montra; quem despende mais de cerca de 3 segundos a observar a montra e vem da esquerda (passa primeiro pela montra e só depois pela porta da loja); quem despende mais de cerca de 11 segundos a observar a montra, vem da direita (tem de voltar atrás para entrar na loja), pára para a observar e varre a montra.

Com base nos resultados obtidos podemos assim distinguir dois tipos de transeuntes: aqueles que entram na loja, sem manifestarem interesse pela montra e os assumem um determinado comportamento face à montra e com base nesse comportamento entram ou não na loja. No primeiro caso, como foi referido anteriormente estamos provavelmente perante potenciais consumidores para os quais a entrada na loja é um comportamento que foi planeado. No segundo caso, existe um interesse que é demonstrado pela montra e, neste caso, o comportamento de entrar parece resultar do poder de atracção da montra em captar a atenção dos indivíduos.

De facto as variáveis usadas nos dois dias para classificar se um indivíduo/grupo entra ou não na loja referem-se ao comportamento que assumem frente à montra, como é visível na Tabela 10.

Nível	Dia 9 de Setembro	Dia 16 de Setembro
1º	Tempo de observação da montra	Tempo de observação da montra
2º	Movimento face à montra	Comportamento Corredor
3º	Pára ou não em frente da montra Tempo de observação da montra	Pára ou não em frente da montra Tempo de observação da montra
4º	Ataque à montra e Pára ou não em frente da montra	Observação Montra
5º	Comportamento face à Montra Deslocação no corredor	Tempo de observação da montra
6º	Sexo Grupo Média Idade	-----
7º	Comportamento Corredor Média Idade	-----
8º	Tempo Observação Montra	-----

Tabela 10: Variáveis de cada nível de cada uma das Árvores de Decisão “Entra ou Não na Loja” do dia 9 e 16 de Setembro

Como se pode constatar, no modelo obtido no dia 9, até ao quinto nível (inclusive), as variáveis consideradas estão relacionadas com o comportamento face à montra, descrevendo o tempo de observação da montra, se permanece quieto em frente dela ou se movimenta-se ao longo da montra, se pára ou não para a observar, se aproxima-se da montra, em que parte da montra mais se fixa para a observar e onde circula no corredor, (se circula pelo interior ou pelo exterior do mesmo). Só após caracterizar o comportamento do consumidor com base nestas variáveis é que o modelo recorre a variáveis relacionadas com o sexo e a idade do potencial consumidor.

No modelo obtido no dia 16 de Setembro todas as variáveis consideradas estão relacionadas com o comportamento face à montra, descrevendo o tempo de observação da montra, se vem da direita ou da esquerda do corredor, se pára ou não para a observar, se foca um dado aspecto da montra ou a varre com o olhar.

Interessa referir que o comportamento de entrar ou não numa loja é resultado da conjugação de diversos factores que tornam difícil prever o comportamento dos consumidores. De facto, na decisão de entrar na loja está subjacente numa primeira instância a decisão da escolha de um dado centro comercial. Esta escolha é fruto de diversos factores: popularidade do centro comercial, ambiente, afluência, percepção que os indivíduos têm do mesmo. Está também subjacente à decisão de entrar na loja, a

decisão de qual o percurso a percorrer para que passe na loja. As lojas situadas em cruzamentos têm uma situação preferencial na medida em que são pontos de passagem de um maior número de pessoas e mais facilmente memorizáveis. A loja objecto de estudo situava-se num cruzamento, pelo que beneficiava desta vantagem competitiva. Por fim, a decisão de entrar é fruto do comportamento e características pessoais de cada indivíduo, da interacção social e da envolvente física e situacional, que contribuem para a decisão tomada com entre 15-30%, 30-50% e 20-24%, respectivamente (Argyle 1976) [2].

Face a este conjunto de factores que determinam o comportamento do consumidor, a envolvente física e situacional é algo em que o lojista pode intervir e influenciar, através nomeadamente, da montra. Esta pode assim revelar-se um factor chave na captação dos clientes, influenciando a entrada na loja. De facto, existe um grupo de indivíduos para o qual a montra é importante, uma vez que a decisão de entrar na loja parece ser motivada pela mesma, dado o comportamento que assume em frente dela, e é nestes consumidores que a loja se deve focalizar levando a cabo os esforços necessários para que o comportamento assumido seja a entrada na loja. De acordo com os resultados obtidos, o factor chave é a montra conseguir captar a atenção do consumidor de modo a que ele se sinta motivado a entrar na loja.

7.1.2. QUAIS AS VARIÁVEIS MAIS IMPORTANTES NA DECISÃO DE COMPRA?

Para classificar quem compra, a primeira variável usada é o tempo na loja, nos dois dias observados, classificando que o indivíduo não compra se esse tempo for inferior a cerca de 11 e 8 minutos, respectivamente. De facto, é de esperar que haja uma relação directa entre a compra e o tempo dispendido na loja, uma vez que é de esperar que quem compra, experimente, logo demore mais tempo na loja.

A segunda variável usada para classificar os indivíduos que permanecem na loja mais tempo que o anteriormente referido é o Sexo do indivíduo/grupo no dia 9 de Setembro e

o Número de indivíduos do grupo no dia 16 de Setembro. Assim no dia 9 de Setembro o maior número de casos classificados como “compra” regista-se para os grupos constituídos pelos dois sexos, que permanecem na loja mais de cerca de 16 minutos. No dia 16 de Setembro, o maior número de casos classificados como “compra” registam-se para as observações compostas por um ou dois indivíduos.

No dia 9 de Setembro compra: os grupos de indivíduos compostos pelos dois sexos, que permanecem na loja mais de cerca de 16 minutos; os homens que passam em frente da montra, que são observados sozinhos e que permanecem na loja entre cerca de 11 e 16 minutos; ou então que são observados em grupo e permanecem na loja mais de cerca de 16 minutos; as mulheres que são observadas em grupo, que permanecem na loja mais de cerca de 11 minutos, passam em frente da montra, vêm da direita ou vêm da esquerda e não olham para a montra, se não passarem em frente da montra, entram se o grupo for constituído por mais de duas mulheres.

No dia 16 de Setembro compram: os grupos compostos por 3 ou mais indivíduos do sexo feminino, que permanecem na loja mais de cerca de 8 minutos, ou pelos dois sexos, que não param em frente da montra e permanecem na loja mais de cerca de 23 minutos; as observações compostas por um ou dois indivíduos, que param em frente da montra, e permanecem na loja mais de cerca de 8 minutos, ou que não param em frente da montra e: deslocam-se no lado da montra e permanecem na loja mais de cerca de 20 minutos ou deslocam-se no lado oposto da montra e permanecem na loja entre cerca de 8 e 10 minutos; as observações compostas por um ou dois indivíduos do sexo feminino ou masculino, que não param em frente da montra, deslocam-se no lado oposto da montra, vêm da esquerda, e permanecem na loja entre cerca de 8 e 10 minutos, e entre 12 e 16 minutos, respectivamente; as observações compostas por dois indivíduos, um de cada sexo, que não param em frente da montra, deslocam-se no lado oposto da montra, vêm da esquerda, olham para a montra e permanecem na loja mais de cerca de 14 minutos.

Podemos assim verificar, conforme está resumido na Tabela 11, que as variáveis usadas nos dois dias para classificar se um indivíduo/grupo compra ou não, referem-se ao tempo que permanecem na loja, ao sexo e número de indivíduos que compõe o grupo e

só posteriormente são consideradas as variáveis relacionadas com o comportamento face à montra.

Nível	Dia 9 de Setembro	Dia 16 de Setembro
1º	Tempo na Loja	Tempo na Loja
2º	Sexo Grupo	Número Indivíduos
3º	Passa ou não em frente montra Tempo na Loja	Pára ou não em frente da montra Sexo Grupo
4º	Comportamento Corredor Número Indivíduos	Pára ou não em frente da montra Deslocação Corredor
5º	Olha ou Não para a Montra Tempo na Loja	Tempo na Loja Comportamento Corredor Tempo Observação
6º	Número Indivíduos	Tempo na Loja Movimento face à montra Número Indivíduos Sexo Grupo
7º	-----	Tempo na Loja Média Idade Olha ou Não para a Montra
8º	-----	Tempo na Loja Média Idade Olha ou Não para a Montra

Tabela 11: Variáveis de cada nível de cada uma das Árvores de Decisão “Compra ou Não” do dia 9 e 16 de Setembro

De facto, no dia 9 de Setembro a compra é função do tempo na loja e do sexo do indivíduo/grupo, só depois é considerado se o consumidor que compra veio da direita ou da esquerda do corredor e se olhou ou não para a montra. No dia 16 de Setembro a compra é função do tempo na loja, do número de indivíduos que compõem cada caso, do sexo do indivíduo/grupo, e só posteriormente se considera se o indivíduo que compra pára ou não em frente da montra, deslocava-se no interior ou exterior do corredor, veio da direita ou da esquerda do corredor, quanto tempo observou a montra e se realizou ou não movimentos em frente da montra.

Assim, a maioria dos indivíduos que compram, não param para observar a montra, pelo que o comportamento face à montra parece por si só não ser decisivo na decisão de comprar. A avaliar pelos resultados obtidos, a compra verifica-se mais quando os indivíduos estão acompanhados.

Está aqui presente a relevância do contexto social em que o indivíduo se encontra, isto é, se está sozinho ou em grupo, e qual o sexo predominante no grupo observado. Só após considerar estas variáveis é que os modelos obtidos têm em conta as que descrevem o comportamento face à montra. De facto existem implicações se a observação for composta por um grupo que resulta dos seguintes factores: pressão que algum elemento do grupo pode fazer para a compra de um determinado artigo que ele próprio gostaria de comprar, o indivíduo que realiza a compra pode ser motivado a escolher aquilo que julga que os restantes elementos gostariam de ter para se sentir admirado.

Um outro elemento muito importante que seria interessante analisar tem a ver com o ambiente da loja: a apresentação dos produtos no interior da loja pode suscitar determinados estados de espírito e disposições que incitem a compra e a forma como os produtos estão disposto pode levar à compra de determinados artigos. Esta é uma informação que não dispomos e que seria interessante analisar em que medida poderia ou não revelar influência na decisão de compra.

Conclui-se assim que não parece existir uma relação directa entre o comportamento face á montra e a compra, uma vez que esta depende mais da composição do grupo que está a ser observado, em termos de sexo e idade.

7.1.3. A DECISÃO DE ENTRAR E DE COMPRAR É OU NÃO INFLUENCIADA POR VARIÁVEIS DIFERENTES CONSOANTE ESTAMOS OU NÃO NA ÉPOCA DE SALDOS?

O que se pretende analisar o eventual efeito da época de saldos, na decisão de entrar na loja e de comprar o produto, uma vez que é um método de promoção de vendas cada vez mais utilizado.

O comportamento face à montra assume maior relevância no dia 16 de Setembro para classificar que um indivíduo entra na loja, uma vez que nesse dia existe uma maior

proporção de casos classificados como entra na loja em função desse comportamento. A época de saldo parece ser por si só um incentivo à entrada na loja, e a montra parece ser um instrumento de atracção e captação de potenciais clientes mais relevante quando está a ser apresentada a nova colecção, isto é, quando é um veículo de informação para os consumidores.

O tempo que o indivíduo permanece na loja, o número de indivíduos que compõem o grupo e o sexo dos mesmos, são as variáveis consideradas mais relevantes na decisão de comprar, no dia 9 de Setembro. Para avaliar o comportamento de quem compra em relação à montra neste dia, apenas são utilizadas as variáveis que permitem constatar se o grupo passa ou não em frente da montra, se vem da direita ou da esquerda do corredor e se olha ou não para a montra. No dia 16 de Setembro, para avaliar esse comportamento, para além das variáveis anteriores, são também usadas as variáveis pára ou não em frente da montra e movimento em frente da montra, pelo que subentende-se que neste caso a montra tem um maior impacto na decisão de compra dos indivíduos.

As variáveis sociais (sexo, número indivíduos grupo, idade) demonstram mais relevância do que as situacionais (impacto da montra) na decisão de compra, pelo que o comportamento face à montra parece demonstrar uma maior influência nessa decisão no dia 16 de Setembro, pois são consideradas mais variáveis que permitem descrever esse comportamento e são classificados mais casos em função dessas variáveis.

Os consumidores demonstraram assim maior interesse pelos artigos expostos na montra da nova colecção Outono-Inverno, pelo que a montra revela-se assim um instrumento mais eficaz de captação de clientes na época da nova colecção.

7.2. PRINCIPAIS CONTRIBUIÇÕES

O presente trabalho pretendeu analisar em que medida o comportamento de um indivíduo/grupo face à montra pode estar relacionado com a entrada na loja e a compra.

Assim, pretendeu-se aferir do poder da montra enquanto instrumento de captação de clientes ao dispor das empresas, de modo a poderem tomar iniciativas que permitam tirar o máximo partido deste instrumento. De facto, esta é uma área de estudo que não tem sido muito explorada e acerca da qual não existem muitos desenvolvimentos pelo que se pretendeu dar alguns passos neste campo, dada a importância que tem, pois a descoberta do que leva o consumidor a entrar e a comprar permitirá tomar medidas no sentido de aumentar o poder de atracção da loja sobre consumidores, bem como o volume de vendas.

Para tal colocou-se a Extração de Conhecimento de Dados (ECD), mais precisamente a classificação, ao dispor deste objectivo, isto é, através do método de indução das Árvores de Decisão obtiveram-se modelos de classificação que permitem, face a novos casos, prever se o indivíduo/grupo entra na loja e se compra. Através deste método pode-se averiguar quais as variáveis relevantes para essa classificação e assim saber em que medida as variáveis relacionadas com o comportamento em relação à montra são importantes.

Concluiu-se que apesar de haver um número considerável de casos em que os indivíduos entram na loja sem olharem para a montra, provavelmente porque já vêm decididos a o fazer, a montra revela influência sobre aqueles que param para a observar, verificando-se uma relação positiva entre o tempo de observação da montra e a entrada na loja. O mesmo acontece no caso da compra, embora aqui exista uma maior relação entre o sexo e o número de indivíduos que compõem cada caso e a decisão de compra do que com as variáveis que descrevem o comportamento face à montra.

Nas duas situações, a época em que é apresentada a nova colecção é aquela onde as variáveis que descrevem o comportamento face à montra estão mais associadas à decisão de entrar e comprar, pelo que concluiu-se que a montra é um instrumento eficaz de captação de cliente, principalmente quando se está a apresentar novos produtos.

Assim, a montra deve ser alvo da atenção dos gestores, na medida em que é eficaz a promover o produto, e é um meio de comunicação fácil de usar e de implementar, que deve ser bem gerido. Uma empresa que consiga usar eficazmente a montra enquanto instrumento de captação dos clientes e veículo de informação terá uma vantagem

competitiva face às concorrentes que se traduzirá num maior volume de vendas e num maior sucesso da loja.

Segundo os resultados obtidos a montra deverá ser capaz de atrair a atenção dos consumidores, na medida em que existe uma relação positiva entre o interesse manifestado pela montra e o comportamento de entrar e de comprar. Esta relação é principalmente visível no caso dos indivíduos que vem da direita e após despenderem algum tempo na visualização da montra voltam atrás e entram na loja. Assim a montra é um importante instrumento de promoção dos produtos, que usada eficazmente se poderá traduzir num aumento das vendas.

Com este trabalho aplicou-se a Análise de Dados a uma área ainda não muito explorada, tendo-se contribuído para o melhor conhecimento da montra e do seu potencial efeito sobre os consumidores.

7.3. LIMITAÇÕES E TRABALHOS FUTUROS

A presente análise permitiu-nos concluir que quer nos períodos pré-saldos e pós-saldos, se o indivíduo/grupo não despender tempo a observar a montra, ele entra na loja em 39,84% e 51,39% dos casos. Dado o elevado número de casos registados nesta situação seria interessante analisar o que motivou em cada observação a entrada na loja: terá sido alguma campanha de marketing ou de publicidade? Será o poder de atracção e o reconhecimento da marca? Esta informação seria útil para se poder actuar e atrair cada vez mais potenciais consumidores.

Assim, uma das limitações que podemos enunciar é o facto de o método de recolha de informação, não nos permitir conhecer os interesses e intenções dos consumidores mas apenas descrever o seu comportamento face à montra.

Por outro lado, apenas se pode efectuar esta análise com uma amostra pois é a única que existe disponível, pelo que um trabalho futuro que seria interessante realizar será a comparação dos modelos obtidos com amostras de vários dias. De facto, esta análise

permitiria confirmar a consistência dos modelos obtidos e também analisar em que medida as conclusões podem ou não ser diferentes consoante o dia da semana ou do mês (o sábado, que é o dia da semana em que se realizou as duas recolhas de dados, é um dos dias de maior afluência aos shoppings, e conseqüentemente de maior volume de vendas). Uma abordagem alternativa seria a de utilizar re-amostras “bootstrap” do conjunto de dados disponível para gerar Árvores de Decisão alternativas e assim estimar a robustez dos resultados. Esse método foi explorado preliminarmente, mas devido a limitações temporais não foram obtidos resultados conclusivos.

Um outro aspecto que seria interessante analisar, acerca do qual não temos informação nos dados recolhidos, tem a ver com a disposição nos artigos no interior da loja, com o layout e o ambiente da mesma (iluminação, musica, etc.), uma vez que estes factores podem ser susceptíveis de alterar a decisão de compra.

8. BIBLIOGRAFIA

- [1] Reynolds, Jonathan and Cuthbertson, Christine (2003), “Retail Strategy, The View from the Bridge”, Butterworth-Heinemann;
- [2] Foxall, Gordan R., Goldsmith, Ronald E. and Brown Stephen (1994), “Consumer Psychology for Marketing”, Routledge;
- [3] McGoldrick, Peter J. (1990), “Retail Marketing”, McGraw-Hill Book Company, Second Edition;
- [4] Aaker, David (1991), “Managing Brand Equity - Capitalizing on the value of a brand name”, New York: The Free Press;
- [5] Aaker, David (1997), “Building brands worthy of devotion”, *Leader to leader*, vol. 11, Winter, pp. 39-42;
- [6] Demetresco, Sylvia (2004), “Como Fazer Montras”, Porto, Vitrimagem.
- [7] Tanaka, Hiroyuki, “A Study for Agent-based Modeling of Migration Behavior of Shoppers”, Center for Spatial Information Sciences of University of Tokyo;
- [8] Dellaert, Benedict G.C., Arentze, Theo A., Timmermans, Harry J.P. (2008), “Shopping context and consumers mental representation of complex shopping trip decision problems”, *Journal of Retailing*, vol 84, n.º 2, pp.219-232;
- [9] Brown, Stephen (1991), “Shopper Circulation in a Planned Shopping Centre”, *Internacionmal Journal of Retail & Distribution Management*, vol 19, n.º 1.
- [10] Hye-Young, Kim, Youn-Kyung, Kim (2008), “Shopping enjoyment and store shopping modes: The moderating influence of chronic time pressure”, *International Journal of Retailing and Consumer Services*, vol. 15, pp. 410-419;

- [11] Richard B. Chase, F. Roberts Jacobs, Nicholas J. Aquilano (2006), “Operations Management for Competitive Advantage with Global Cases”, McGraw-Hill International Edition;
- [12] Blattberg, R. and S. Neslin (1990), “Sales Promotions: Concepts, Methods and Strategies”, Englewood Cliffs, New Jersey: Prentice Hall;
- [13] Belch, George E. e Belch, Michael A. (2001), “Advertising and Promotion – An Integrated Marketing Communications Perspective”, Fifth Edition, McGraw-Hill.
- [14] DelVecchio, Devon, Henard, David H., Freling Traci H. (2006), “The effect of sales promotion on post-promotion brand preference: A meta-analysis”, Journal of Retailing, vol 82, n.º 3, pp, 203-213;
- [15] Brito, Pedro Quelhas e Hammond, Kathy (2007), “Strategic Versus Tactical Nature of Sales Promotions”, Jornal of Marketing Communications, Vol 13, n.º 2, pp. 131-148;
- [16] Witten, Ian H. e Frank, Eibe (2005), “Data Mining - Practical Machine Learning Tools and Techniques”, Second Edition, Morgan Kaufmann;
- [17] Han, Jiawei e Kamber, Micheline (2006), “Data Mining – Concepts and Techniques”, Second Edition, Morgan Kaufmann;
- [18] J. Ross (1993), “C4.5: programs for machine learning”, Morgan Kaufmann;
- [19] Witten, Ian H. e Frank, Eibe; (2000); “Weka: Practical Machine Learning Tools and Techniques with Java Implementations”, working paper of the Computing and Mathematical Sciences Papers;
- [20] Bouckaert, Remco R.; Frank, Eibe; Hall, Mark; Kirkby, Richard; Reutemann, Peter; Seewald , Alex; Scuse, David (2009), “Weka Manual for version 3-7-0”, The University of Waikato, Hamilton, New Zealand;

- [21] Ribeiro, António (2008), “Impacto da Comunicação Visual Veiculada pela Montra da Loja no Comportamento do Transeunte”, Tese de Mestrado em Marketing na Faculdade de Economia do Porto;
- [22] Brown, James R. E Dant, Rajiv P. (2008), “Scientific method and retailing research: A retrospective”, *Journal of Retailing*, vol 84, n.º 1, pp. 1-13;
- [23] Bento, José Ferreira Murteira (1993), “Análise Exploratória de Dados”; Mc Gaw Hill;

9. ANEXOS

ANEXO N.º 1: COMPORTAMENTO FACE À MONTRA

Não Passa Vs Passa Frente da Montra	Não Olha Vs Olha Montra	Ataque Montra	Movimento Montra	Não Pára Vs Pára	Observação Montra	Comportamento Montra	Dia 9 de Setembro			Dia 16 de Setembro			Total
							Não Compra	Não Compra	Entrar Compra	Não Compra	Não Compra	Entrar Compra	
Não Passa em Frente da Montra	Não Olha para a Montra	Sem Ataque	Sem Movimento	Não Pára	Sem Atitude	Sem Comportamento			294			366	70
Passa em Frente da Montra	Não Olha para a Montra	Sem Ataque	Sem Movimento	Não Pára	Sem Atitude	Sem Comportamento			134			195	45
		Ataque Natural	Quieto	Não Pára	Quieto	Fixo à Direita Fixo ao Centro						1	1
	Olha para a Montra		Com Movimento	Pára	Movimento	Combinado Combinado	729	6	3	517	5		1.260
							98	7	1	125	21	1	253
							6	12	2	17	10	3	50
							5	7		12	5	2	31
							7	1		8	3		19
	Ataque de Aproximação		Quieto	Pára	Movimento	Fixo à Esquerda Fixo ao Centro						1	1
			Com Movimento	Não Pára	Movimento	Combinado	96	2		94	1	1	194
							17	3		22	6		48
							1.121	742	135	912	968	171	4.049

**ANEXO N.º 2: OUTPUT GERADO PELO WEKA PARA O MODELO DAS ÁRVORES DE
DECISÃO EM RELAÇÃO AO ATRIBUTO “ENTRA OU NÃO NA LOJA” NO DIA 9 DE
SETEMBRO**

```

=== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    importancia_montra
Instances:   1998
Attributes:  14
              ComportamentoCorredor
              DeslocacaoCorredor
              NaoPassaVsPassaFrenteMontra
              NaoOlhaVsOlhaMontra
              AtaqueMontra
              NaoPáraVsPára
              MovimentoMontra
              ComportamentoMontra
              ObservacaoMontra
              TempoObservacao(Seg)
              EntraVsNaoEntra
              NúmeroIndividuos
              SexoGrupo
              MediaIdade
Test mode:   10-fold cross-validation

=== Classifier model (full training set) ===

J48 pruned tree
-----

TempoObservacao(Seg) <= 0: 1 (796.0)
TempoObservacao(Seg) > 0
|   MovimentoMontra = 0: 0 (0.0)
|   MovimentoMontra = 1
|   |   NaoPáraVsPára = 0: 1 (3.0)
|   |   NaoPáraVsPára = 1
|   |   |   AtaqueMontra = 0: 0 (0.0)
|   |   |   AtaqueMontra = 1: 0 (197.0/34.0)
|   |   |   AtaqueMontra = 2
|   |   |   |   ComportamentoMontra = 0: 1 (0.0)
|   |   |   |   ComportamentoMontra = 1
|   |   |   |   |   SexoGrupo = F
|   |   |   |   |   |   ComportamentoCorredor = 1: 1 (4.0)
|   |   |   |   |   |   ComportamentoCorredor = 2
|   |   |   |   |   |   |   TempoObservacao(Seg) <= 12: 1 (3.0)
|   |   |   |   |   |   |   TempoObservacao(Seg) > 12: 0 (3.0)
|   |   |   |   |   |   |   SexoGrupo = M: 1 (1.0)
|   |   |   |   |   |   |   SexoGrupo = Misto
|   |   |   |   |   |   |   MediaIdade <= 2: 1 (5.0)
|   |   |   |   |   |   |   MediaIdade > 2: 0 (4.0/1.0)

```

```

|      |      |      |      ComportamentoMontra = 2
|      |      |      |      |      SexoGrupo = F: 1 (2.0)
|      |      |      |      |      SexoGrupo = M: 0 (4.0)
|      |      |      |      |      SexoGrupo = Misto: 1 (6.0/1.0)
|      |      |      |      ComportamentoMontra = 3: 0 (8.0/1.0)
|      |      |      |      ComportamentoMontra = 4: 1 (0.0)
|      MovimentoMontra = 2
|      |      TempoObservacao(Seg) <= 9: 0 (921.0/14.0)
|      |      TempoObservacao(Seg) > 9
|      |      |      NaoPáraVsPára = 0: 0 (29.0)
|      |      |      NaoPáraVsPára = 1
|      |      |      DeslocacaoCorredor = 1: 1 (7.0/1.0)
|      |      |      DeslocacaoCorredor = 2
|      |      |      MediaIdade <= 2.6: 1 (2.0)
|      |      |      MediaIdade > 2.6: 0 (3.0)

Number of Leaves :    22

Size of the tree :    36

Time taken to build model: 0.08 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          1927           96.4464
%
Incorrectly Classified Instances         71           3.5536
%
Kappa statistic                        0.9274
Mean absolute error                    0.0546
Root mean squared error                0.1777
Relative absolute error                 11.0882 %
Root relative squared error            35.7997 %
Total Number of Instances              1998

=== Detailed Accuracy By Class ===

TP Rate    FP Rate    Precision    Recall    F-Measure    ROC Area
Class
    0.991    0.07      0.948      0.991      0.969      0.975    0
    0.93     0.009     0.988      0.93      0.958      0.975    1

=== Confusion Matrix ===

      a      b  <-- classified as
1111    10 |      a = 0
 61   816 |      b = 1

```

**ANEXO N.º 3: OUTPUT GERADO PELO WEKA PARA O MODELO DAS ÁRVORES DE
DECISÃO EM RELAÇÃO AO ATRIBUTO “ENTRA OU NÃO NA LOJA” NO DIA 16 DE
SETEMBRO**

```
=== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    importancia_montra
Instances:   2051
Attributes:  14
              ComportamentoCorredor
              DeslocacaoCorredor
              NaoPassaVsPassaFrenteMontra
              NaoOlhaVsOlhaMontra
              AtaqueMontra
              NaoPáraVsPára
              MovimentoMontra
              ComportamentoMontra
              ObservacaoMontra
              TempoObservacao(Seg)
              EntraVsNaoEntra
              NúmeroIndividuos
              SexoGrupo
              MediaIdade
Test mode:   10-fold cross-validation

=== Classifier model (full training set) ===

J48 pruned tree
-----

TempoObservacao(Seg) <= 0: 1 (1054.0)
TempoObservacao(Seg) > 0
|   ComportamentoCorredor = 1
|   |   NaoPáraVsPára = 0: 0 (597.0/5.0)
|   |   NaoPáraVsPára = 1
|   |   |   ObservacaoMontra = 0: 0 (0.0)
|   |   |   ObservacaoMontra = 1: 0 (174.0/21.0)
|   |   |   ObservacaoMontra = 2
|   |   |   |   TempoObservacao(Seg) <= 11: 0 (141.0/4.0)
|   |   |   |   TempoObservacao(Seg) > 11: 1 (12.0/1.0)
|   |   ComportamentoCorredor = 2
|   |   |   TempoObservacao(Seg) <= 3: 0 (27.0)
|   |   |   TempoObservacao(Seg) > 3: 1 (46.0/2.0)

Number of Leaves   :    8

Size of the tree   :   14

Time taken to build model: 0.09 seconds
```

```

=== Stratified cross-validation ===
=== Summary ===

```

```

Correctly Classified Instances      2016           98.2935
%
Incorrectly Classified Instances    35           1.7065
%
Kappa statistic                    0.9656
Mean absolute error                 0.0305
Root mean squared error             0.1269
Relative absolute error              6.1665 %
Root relative squared error         25.5304 %
Total Number of Instances          2051

```

```

=== Detailed Accuracy By Class ===

```

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	
0.997	0.028	0.966	0.997	0.981	0.99	0
0.972	0.003	0.997	0.972	0.984	0.99	1

```

=== Confusion Matrix ===

```

```

      a      b  <-- classified as
909      3 |      a = 0
32 1107 |      b = 1

```

**ANEXO N.º 4: OUTPUT GERADO PELO WEKA PARA O MODELO DAS ÁRVORES DE
DECISÃO EM RELAÇÃO AO ATRIBUTO “COMPRA OU NÃO” NO DIA 9 DE SETEMBRO**

=== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: importancia_montra
Instances: 877
Attributes: 15

ComportamentoCorredor
DeslocacaoCorredor
NaoPassaVsPassaFrenteMontra
NaoOlhaVsOlhaMontra
AtaqueMontra
NaoParaVsPara
MovimentoMontra
ComportamentoMontra
ObservacaoMontra
TempoObservacao(seg)
TempoLoja(seg)
NúmeroIndividuos
SexoGrupo
MediaIdade
Compra

Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

J48 pruned tree

TempoLoja(seg) <= 633: 0 (703.0/46.0)

TempoLoja(seg) > 633

```
|   SexoGrupo = F
|   |   NaoPassaVsPassaFrenteMontra = 0
|   |   |   NúmeroIndividuos <= 2: 0 (31.0/10.0)
|   |   |   NúmeroIndividuos > 2: 1 (2.0)
|   |   NaoPassaVsPassaFrenteMontra = 1
|   |   |   ComportamentoCorredor = 1: 1 (4.0/1.0)
|   |   |   ComportamentoCorredor = 2
|   |   |   |   NaoOlhaVsOlhaMontra = 0
|   |   |   |   |   NúmeroIndividuos <= 1: 0 (10.0/4.0)
|   |   |   |   |   NúmeroIndividuos > 1: 1 (9.0/3.0)
|   |   |   |   NaoOlhaVsOlhaMontra = 1: 0 (28.0/11.0)
|   SexoGrupo = M
|   |   NaoPassaVsPassaFrenteMontra = 0: 0 (4.0)
|   |   NaoPassaVsPassaFrenteMontra = 1
|   |   |   NúmeroIndividuos <= 1
|   |   |   |   TempoLoja(seg) <= 933: 1 (3.0)
|   |   |   |   TempoLoja(seg) > 933: 0 (4.0/1.0)
|   |   |   NúmeroIndividuos > 1: 1 (5.0)
```

```

|   SexoGrupo = Misto
|   |   TempoLoja(seg) <= 985: 0 (30.0/12.0)
|   |   TempoLoja(seg) > 985: 1 (44.0/12.0)

Number of Leaves   :    13

Size of the tree   :    24

Time taken to build model: 0.11 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          734           83.6944
%
Incorrectly Classified Instances        143           16.3056
%
Kappa statistic                        0.2752
Mean absolute error                    0.2077
Root mean squared error                0.3532
Relative absolute error                79.5571 %
Root relative squared error            97.8615 %
Total Number of Instances              877

=== Detailed Accuracy By Class ===

TP Rate    FP Rate    Precision    Recall    F-Measure    ROC Area
Class
  0.934     0.696     0.881     0.934     0.906     0.665    0
  0.304     0.066     0.456     0.304     0.364     0.665    1

=== Confusion Matrix ===

  a    b  <-- classified as
693  49 |    a = 0
 94  41 |    b = 1

```

**ANEXO N.º 5: OUTPUT GERADO PELO WEKA PARA O MODELO DAS ÁRVORES DE
DECISÃO EM RELAÇÃO AO ATRIBUTO “COMPRA OU NÃO” NO DIA 16 DE SETEMBRO**

=== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: importancia_montra
Instances: 1139
Attributes: 15

ComportamentoCorredor
DeslocacaoCorredor
NaoPassaVsPassaFrenteMontra
NaoOlhaVsOlhaMontra
AtaqueMontra
NaoParaVsPara
MovimentoMontra
ComportamentoMontra
ObservacaoMontra
TempoObservacao(seg)
TempoLoja(seg)
NúmeroIndividuos
SexoGrupo
MediaIdade
Compra

Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

J48 pruned tree

TempoLoja(seg) <= 494: 0 (841.0/34.0)

TempoLoja(seg) > 494

| NúmeroIndividuos <= 2

| | NaoParaVsPara = 0

| | | DeslocacaoCorredor = 1

| | | | TempoLoja(seg) <= 1196

| | | | | MovimentoMontra = 0

| | | | | | MediaIdade <= 2.25: 0 (30.0/11.0)

| | | | | | MediaIdade > 2.25: 1 (6.0/1.0)

| | | | | MovimentoMontra = 1: 0 (24.0/8.0)

| | | | | MovimentoMontra = 2: 0 (1.0)

| | | | TempoLoja(seg) > 1196: 1 (27.0/6.0)

| | | DeslocacaoCorredor = 2

| | | | ComportamentoCorredor = 1

| | | | | TempoLoja(seg) <= 594: 0 (17.0/2.0)

| | | | | TempoLoja(seg) > 594: 1 (54.0/23.0)

| | | | ComportamentoCorredor = 2

| | | | | SexoGrupo = F

| | | | | | TempoLoja(seg) <= 626: 1 (6.0/1.0)

| | | | | | TempoLoja(seg) > 626: 0 (36.0/6.0)

```

| | | | | SexoGrupo = M
| | | | | | TempoLoja(seg) <= 939
| | | | | | | TempoLoja(seg) <= 698: 0 (2.0)
| | | | | | | TempoLoja(seg) > 698: 1 (4.0)
| | | | | | | TempoLoja(seg) > 939: 0 (8.0)
| | | | | SexoGrupo = Misto
| | | | | | TempoLoja(seg) <= 809: 0 (9.0)
| | | | | | TempoLoja(seg) > 809
| | | | | | | NaoOlhaVsOlhaMontra = 0: 0 (8.0/3.0)
| | | | | | | NaoOlhaVsOlhaMontra = 1: 1 (9.0/3.0)
| | | NaoParaVsPara = 1
| | | | DeslocacaoCorredor = 1
| | | | | TempoObservacao(Seg) <= 7: 0 (2.0)
| | | | | TempoObservacao(Seg) > 7: 1 (3.0/1.0)
| | | | DeslocacaoCorredor = 2: 1 (4.0)
| NúmeroIndividuos > 2
| | SexoGrupo = F: 1 (4.0)
| | SexoGrupo = M: 0 (1.0)
| | SexoGrupo = Misto
| | | NaoParaVsPara = 0
| | | | TempoLoja(seg) <= 1353
| | | | | NúmeroIndividuos <= 3
| | | | | | NaoOlhaVsOlhaMontra = 0
| | | | | | | MediaIdade <= 2.75: 1 (8.0)
| | | | | | | MediaIdade > 2.75: 0 (3.0/1.0)
| | | | | | | NaoOlhaVsOlhaMontra = 1: 0 (5.0/1.0)
| | | | | NúmeroIndividuos > 3
| | | | | | TempoLoja(seg) <= 644
| | | | | | | TempoLoja(seg) <= 517: 0 (2.0)
| | | | | | | TempoLoja(seg) > 517: 1 (3.0)
| | | | | | | TempoLoja(seg) > 644: 0 (7.0)
| | | | | | TempoLoja(seg) > 1353: 1 (11.0)
| | | | NaoParaVsPara = 1: 0 (4.0/1.0)

Number of Leaves : 29

Size of the tree : 54

Time taken to build model: 0.08 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          980           86.0404
%
Incorrectly Classified Instances        159           13.9596
%
Kappa statistic                        0.3865
Mean absolute error                    0.1768
Root mean squared error                0.3281
Relative absolute error                69.1451 %
Root relative squared error            91.8433 %

```

Total Number of Instances				1139			
=== Detailed Accuracy By Class ===							
TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area		
Class							
0.941	0.596	0.899	0.941	0.92	0.743	0	
0.404	0.059	0.548	0.404	0.465	0.743	1	
=== Confusion Matrix ===							
a	b	<-- classified as					
911	57		a = 0				
102	69		b = 1				