



**UNIVERSIDADE DE ÉVORA**

**ESCOLA DE CIÊNCIAS E TECNOLOGIA**

DEPARTAMENTO DE MATEMÁTICA

**CARACTERIZAÇÃO DOS HÁBITOS  
DE E-COMMERCE**

**João Filipe de Souta Neiva Machado**

Orientador: Prof.<sup>a</sup> Doutora Anabela Cristina Cavaco Ferreira  
Afonso

Co-Orientador: Prof. Doutor Manuel Joaquim Piteira Minhoto

**Mestrado em Modelação Estatística e Análise de Dados**

Área de especialização: **Modelação Estatística e Análise de Dados**

Dissertação

Évora, 2015



**UNIVERSIDADE DE ÉVORA**

**ESCOLA DE CIÊNCIAS E TECNOLOGIA**

DEPARTAMENTO DE MATEMÁTICA

**CARACTERIZAÇÃO DOS HÁBITOS  
DE E-COMMERCE**

**João Filipe de Souta Neiva Machado**

Orientador: Prof.<sup>a</sup> Doutora Anabela Cristina Cavaco Ferreira  
Afonso

Co-Orientador: Prof. Doutor Manuel Joaquim Piteira Minhoto

**Mestrado em Modelação Estatística e Análise de Dados**

Área de especialização: **Modelação Estatística e Análise de Dados**

Dissertação

Évora, 2015

# Resumo

O fenómeno do *e-commerce* tem vindo a crescer em Portugal e já faz parte do dia-a-dia dos portugueses, complementarmente e/ou em alternativa ao comércio tradicional. Por isso, cada vez mais as empresas procuram compreender as tendências que moldam o comportamento do consumidor *online* e podem influenciar as suas práticas de consumo.

Nesse sentido, este estudo tem como objetivo geral caracterizar os hábitos de *e-commerce* dos portugueses com 18 ou mais anos. Para o efeito foi elaborado um questionário e recolhida uma amostra pelo processo de amostragem em bola de neve. No processo de análise de dados, foram utilizadas as técnicas Análise de Correspondências Múltiplas, função de Fisher resultante da Análise Discriminante e Regressão Logística, também com o objectivo de compreender a coerência entre os resultados e a complementariedade entre elas.

Participaram no estudo 400 indivíduos, sendo a maioria mulheres (58%), residentes em diversos distritos de Portugal, com idades compreendidas entre os 18 e os 71 anos, tendo a maior parte um grau de instrução universitária.

Os resultados permitiram concluir quais os produtos e serviços com maior compra/adesão *online* para cada um dos sexos, assim como inferir sobre as principais vantagens e desvantagens associadas à compra *online*, para ambos os sexos e para compradores e não compradores. No que diz respeito à percepção de segurança nas compras *online*, este estudo revelou que metade dos participantes consideram este modo de compra seguro, sendo os fatores mais importantes para a percepção de segurança a credibilidade da empresa/marca do produto e a variedade de meios de pagamento disponíveis.

**Palavras-Chave:** *E-Commerce*, Consumidor *Online*, Análise de Correspondências Múltiplas, Análise Discriminante, Regressão Logística, Análise de Clusters.

# Abstract

## CHARACTERIZATION OF E-COMMERCE BEHAVIOR

The phenomenon of e-commerce has been growing in Portugal and is already part of the day to day life of the Portuguese, complementary and/or alternatively to traditional trade. Therefore, more and more companies seek to understand the trends shaping the online consumer behavior that can influence their consumption practices.

In this sense, this study has the general objective to characterize the e-commerce habits of the Portuguese with 18 or more years. For this purpose, a questionnaire was prepared and a sample was collected with a snowball sampling process. In the process of data analysis, we used the following techniques: Multiple Correspondence Analysis, Fisher function resulting from the Discriminant Analysis and Logistic Regression. This study also has the aim of understanding the consistency between the results and the complementarity among these techniques.

In this study, participated 400 individuals, the majority being women (58%), resident in several districts of Portugal, aged between 18 and 71 years, with most having a degree of university education.

The results allowed to know the goods and services with higher purchase/online adherence by gender, as well as to inference about the main advantages and disadvantages associated with online shopping, for both sexes and for both buyers and no buyers. With regard to the security perception in online shopping, this study revealed that half of the participants consider online buying safe, and that the most important factors for the security perception are the credibility of the company / brand of the product and the variety of media available payment.

**Keywords:** E-Commerce, Online consumer behavior, Multiple Correspondence Analysis, Discriminant Analysis, Logistic Regression, Clusters analysis.

# Agradecimentos

É com muita satisfação que expresso aqui um profundo agradecimento a todos aqueles que tornaram possível a realização deste trabalho.

Gostaria de agradecer à Professora Doutora Anabela Afonso e ao Professor Doutor Manuel Minhoto, orientadora e co-orientador desta tese, pelo apoio, incentivo e acima de tudo por toda a disponibilidade demonstradas.

A todos os participantes no estudo pela sua colaboração e divulgação do questionário.

À minha mãe, Margarida, por todos os sacrifícios, dedicação e força que deu à minha vida e em particular à minha educação.

A todos os familiares, amigos e colegas a quem não pude dedicar-me no decorrer deste trabalho.

À Andresa, por todo o amor, carinho, compreensão e paciência. E por ser o pilar da minha vida.

Por último, ao Tomás que ainda não nasceu, mas que já ocupa um lugar tão grande na minha vida.

# Conteúdo

1)	INTRODUÇÃO .....	1
1.1)	EVOLUÇÃO DA <i>INTERNET</i> .....	1
1.2)	EVOLUÇÃO DO <i>E-COMMERCE</i> .....	5
1.3)	MODELOS DE COMPORTAMENTO DO CONSUMIDOR <i>ONLINE</i> .....	7
1.4)	PERFIL SOCIODEMOGRÁFICO DO CONSUMIDOR <i>ONLINE</i> .....	8
1.6)	PERCEÇÕES DE RISCO ASSOCIADAS AO <i>E-COMMERCE</i> .....	11
1.7)	MEIOS DE PUBLICIDADE <i>ONLINE</i> .....	12
1.8)	ESTUDOS REALIZADOS SOBRE <i>E-COMMERCE</i> EM PORTUGAL .....	13
1.9)	PERTINÊNCIA E OBJETIVOS DO ESTUDO .....	14
2)	MÉTODOS ESTATÍSTICOS .....	16
2.1)	COEFICIENTES DE CORRELAÇÃO .....	16
2.2)	ANÁLISE DE CORRESPONDÊNCIAS MÚLTIPLAS.....	18
2.3)	ANÁLISE DISCRIMINANTE .....	23
2.4)	REGRESSÃO LOGÍSTICA.....	24
3)	METODOLOGIA .....	26
3.1)	HIPÓTESES TEÓRICAS EM ESTUDO .....	26
3.2)	CARACTERIZAÇÃO DO UNIVERSO DO ESTUDO E DEFINIÇÃO DA AMOSTRA .....	27
3.3)	CONSTRUÇÃO DO QUESTIONÁRIO E RECOLHA DA INFORMAÇÃO .....	28
4)	APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS .....	30
4.1)	ANÁLISE SOCIODEMOGRÁFICA DA AMOSTRA .....	30
4.2)	EXPERIÊNCIA <i>ONLINE</i> DOS INQUIRIDOS .....	33
4.3)	COMPRADORES <i>ONLINE</i> .....	35
4.4)	NÃO COMPRADORES <i>ONLINE</i> .....	62
4.5)	SEGURANÇA NAS COMPRAS <i>ONLINE</i> .....	63
4.6)	PREFERÊNCIAS .....	66
4.6.1)	VANTAGENS DA COMPRA <i>ONLINE</i> .....	66
4.6.2)	DESVANTAGENS DA COMPRA <i>ONLINE</i> .....	83
4.6.3)	MOTIVAÇÕES E PREFERÊNCIAS NA COMPRA <i>ONLINE</i> .....	101
4.6.4)	MEIOS DE COMUNICAÇÃO <i>ONLINE</i> .....	103
5)	CONSIDERAÇÕES FINAIS .....	122
5.1)	LIMITAÇÕES DO ESTUDO.....	122
5.2)	CONCLUSÕES SOBRE AS HIPÓTESES COLOCADAS .....	122
	BIBLIOGRAFIA.....	125
	ANEXOS .....	128

# Lista de Tabelas

Tabela 1: Principais lançamentos da história da <i>internet</i> nos primeiros anos do século XXI.....	4
Tabela 2: Tabela de contingência.....	16
Tabela 3: Poder discriminativo do modelo de regressão logística com base na área abaixo da curva ROC... ..	25
Tabela 4: Hábito de comprar <i>online</i> por sexo .....	34
Tabela 5: Medidas de discriminação associadas aos produtos comprados <i>online</i> .....	37
Tabela 6: Posicionamento das variáveis mais relevantes para a ACM dos produtos comprados <i>online</i> .....	39
Tabela 7: Caracterização dos 2 grupos identificados, relativamente aos às vantagens da compra <i>online</i> identificadas pelos inquiridos.....	43
Tabela 8: Modelo de Análise Discriminante inicialmente obtido para os produtos comprados <i>Online</i> , usando como variável de grupo o “Sexo” .....	44
Tabela 9: Matriz de estrutura da Análise Discriminante inicialmente obtido para os produtos comprados <i>Online</i> , usando como variável de grupo o “Sexo” .....	45
Tabela 10: Modelo final de Análise Discriminante obtido para os produtos comprados <i>Online</i> , usando como Variável de grupo “Sexo” .....	45
Tabela 11: Coeficientes dos produtos comprados <i>online</i> na função de Fisher por sexo .....	46
Tabela 12: Percentagem de homens e mulheres classificados corretamente, através do modelo obtido para os produtos comprados <i>online</i> .....	47
Tabela 13: Modelo inicial de Regressão Logística obtido para o sexo em função dos produtos comprados <i>online</i> .....	47
Tabela 14: Modelo de Regressão Logística obtido para o sexo em função dos produtos comprados <i>online</i> .....	48
Tabela 15: Valores previstos pelo modelo de regressão logística para o sexo em função dos produtos comprados <i>online</i> , usando o ponto de corte 0,50 .....	48
Tabela 16: Tabela resumo dos métodos utilizados na análise dos produtos comprados <i>Online</i> .....	50
Tabela 17: Medidas de discriminação associadas à ACM dos serviços aderidos <i>online</i> .....	52
Tabela 18: Posicionamento das variáveis mais relevantes para a ACM dos serviços aderidos <i>online</i> .....	53
Tabela 19: Caracterização dos grupos identificados, relativamente aos aos serviços aderidos <i>online</i> .....	55
Tabela 20: Modelo final de Análise Discriminante obtido para os serviços aderidos <i>online</i> , usando como variável de grupo “Sexo” .....	56
Tabela 21: Coeficientes da função de Fisher associada à Análise Discriminante os serviços aderidos <i>Online</i> , usando como variável de grupo “Sexo” .....	56
Tabela 22: Percentagem de homens e mulheres classificados corretamente, através do modelo obtido para os serviços aderidos <i>online</i> .....	57
Tabela 23: Modelo de Regressão Logística final obtido para o sexo em função dos serviços aderidos <i>online</i> .....	57
Tabela 24: Valores previstos pelo modelo de regressão logística para o sexo em função dos serviços aderidos <i>online</i> , usando o ponto de corte 0,50 .....	58
Tabela 25: Tabela resumo dos métodos utilizados na análise dos serviços aderidos <i>Online</i> .....	59
Tabela 26: Hábitos de compra Online - para o próprio ou para outros .....	59
Tabela 27: Frequência das compras online por sexo .....	60
<b>Tabela 28:</b> Programação prévia das compras por sexo.....	61
Tabela 29: Medidas de discriminação associadas à ACM das vantagens da compra <i>Online</i> .....	70
Tabela 30: Posicionamento das variáveis mais relevantes para a ACM das vantagens de compra <i>online</i> .....	71

Tabela 31: Caracterização dos 5 grupos identificados, relativamente aos às vantagens da compra <i>online</i> identificadas pelos inquiridos.....	74
Tabela 32: Modelo de Análise Discriminante para as vantagens na compra <i>Online</i> , usando como variável de agrupamento “Costuma realizar compras <i>online</i> ” .....	75
Tabela 33: Coeficientes da função de Fisher associada à Análise Discriminante para as vantagens da compra <i>Online</i> , usando como variável de agrupamento “Costuma realizar compras <i>online</i> ” .....	75
Tabela 34: Percentagem de compradores e não compradores classificados corretamente, através do modelo obtido para as vantagens da compra <i>online</i> .....	76
<b>Tabela 35:</b> Modelo de Análise Discriminante para as vantagens na compra <i>Online</i> , usando como variável de agrupamento “Sexo” .....	77
Tabela 36: Coeficientes da função de Fisher associada à Análise Discriminante para as vantagens da compra <i>Online</i> , usando como variável de agrupamento “Sexo” .....	77
<b>Tabela 37:</b> Percentagem de homens e mulheres classificados corretamente, através do modelo obtido às vantagens das compras <i>online</i> .....	78
Tabela 38: Modelo de Regressão Logística para os compradores online em função das vantagens na compra online .....	78
Tabela 39: Valores previstos pelo modelo de regressão logística para os compradores em função das vantagens associadas à compra <i>online</i> , usando o ponto de corte 0,70 .....	79
Tabela 40: Modelo de Regressão Logística para o sexo em função das vantagens .....	80
Tabela 41: Valores previstos pelo modelo de regressão logística para o sexo em função das vantagens associadas às compras <i>online</i> , usando o ponto de corte 0,50 .....	80
Tabela 42: Tabela resumo dos métodos utilizados na análise das vantagens em comprar <i>Online</i> .....	82
Tabela 43: Medidas de discriminação associadas à ACM das desvantagens da compra <i>online</i> .....	88
Tabela 44: Posicionamento das variáveis mais relevantes para a ACM das desvantagens de compra <i>online</i> .....	89
Tabela 45: Caracterização dos 5 grupos identificados, relativamente aos às desvantagens da compra <i>online</i> identificadas pelos inquiridos.....	92
Tabela 46: Modelo de Análise Discriminante para as desvantagens na compra <i>Online</i> , usando como Variável de grupo “Costuma realizar compras <i>online</i> ” .....	93
Tabela 47: Coeficientes da função de Fisher associada à Análise Discriminante para as desvantagens da compra <i>Online</i> , usando como variável de grupo “Costuma realizar compras <i>online</i> ” .....	93
Tabela 48: Probabilidade de um novo elemento pertencer ao grupo dos compradores, através do modelo obtido para as desvantagens da compra online .....	94
Tabela 49: Modelo de Análise Discriminante para as desvantagens na compra <i>Online</i> , usando como variável de grupo “Sexo” .....	94
Tabela 50: Coeficientes da função de Fisher associada à Análise Discriminante para as desvantagens da compra <i>Online</i> , usando como variável de grupo “Sexo” .....	95
Tabela 51: Percentagem de homens e mulheres classificados corretamente, através do modelo obtido para as desvantagens associadas às compras <i>online</i> .....	96
Tabela 52: Modelo de Regressão Logística para os compradores <i>online</i> em função das desvantagens na compra <i>online</i> .....	96
Tabela 53: Valores previstos com o Modelo de Regressão Logística obtido para os compradores em função das desvantagens, usando o ponto de corte 0,70.....	97
Tabela 54: Modelo de Regressão Logística obtido para as desvantagens na compra <i>Online</i> , usando como variável suplementar “Costuma realizar compras <i>online</i> ” .....	98



Tabela 55: Valores previstos com o Modelo de Regressão Logística obtido para os sexos em função das desvantagens, usando o ponto de corte 0,40 .....	99
Tabela 56: Tabela resumo dos métodos utilizados na análise das desvantagens em comprar <i>Online</i> .....	100
Tabela 57: Quais os motivos que o poderiam levar a comprar/comprar mais <i>online</i> ?.....	101
Tabela 58: Medidas de discriminação associadas à ACM dos meios de publicidade mais eficazes.....	108
Tabela 59: Posicionamento das variáveis mais relevantes para a ACM dos meios de publicidade mais eficazes.....	109
Tabela 60: Caracterização dos 3 grupos identificados, relativamente aos meios de publicidade mais eficazes, identificados pelos inquiridos. ....	112
Tabela 61: Modelo final de Análise Discriminante obtido para os meios de publicidade mais eficazes, variável de grupo “Costuma realizar compras <i>online</i> ” .....	113
Tabela 62: Coeficientes da função de Fisher associada à Análise Discriminante obtido para os meios de publicidade mais eficazes, variável de grupo “Costuma realizar compras <i>online</i> ” .....	113
Tabela 63: Probabilidade de um novo elemento pertencer ao grupo dos compradores, através do modelo obtido para os meios de publicidade mais eficazes .....	114
Tabela 64: Modelo final de Análise Discriminante obtido para os meios de publicidade mais eficazes, variável de grupo o sexo .....	114
Tabela 65: Coeficientes da função de Fisher associada à Análise Discriminante para os meios de publicidade mais eficazes, usando como Variável de grupo “Sexo” .....	115
Tabela 66: Percentagem de homens e mulheres classificados corretamente, através do modelo obtido para os meios de publicidade mais eficazes.....	116
Tabela 67: Modelo de Regressão Logística obtido para os meios de <i>publicidade mais eficazes</i> , usando como variável suplementar “Costuma realizar compras <i>online</i> ” .....	116
Tabela 68: Valores previstos com o Modelo de Regressão Logística para os compradores online em função dos meios de publicidade <i>online</i> , usando o ponto de corte 0,70.....	117
Tabela 69: Modelo de Regressão Logística obtido para os meios publicitários mais eficazes, usando como variável suplementar “Costuma realizar compras <i>online</i> ” .....	118
Tabela 70: Valores previstos com o Modelo de Regressão Logística para os sexos em função dos meios de publicidade <i>online</i> , usando o ponto de corte 0,40.....	119
Tabela 71: Tabela resumo dos métodos utilizados na análise dos meios de publicidade mais eficazes em comprar <i>Online</i> .....	121
Tabela 72: Resumo dos resultados para as hipóteses do estudo.....	123

# Lista de Figuras

Figura 1: Mapa com evolução das primeiras redes ( <i>nets</i> ).....	2
Figura 2: Fluxo do questionário aplicado .....	29
Figura 3: Percentagem de individuos da amostra divididos pelos distritos de Portugal.....	30

# Lista de Gráficos

Gráfico 1: Utilizadores de <i>internet</i> por continente .....	3
Gráfico 2: Penetração de Internet na População (Junho/2014).....	4
<b>Gráfico 3:</b> Sexo dos inquiridos.....	31
Gráfico 4: Caixa de bigodes idade por sexo.....	31
<b>Gráfico 5:</b> Distribuição dos indivíduos nas classes de idade, por sexo.....	32
<b>Gráfico 6:</b> Distribuição dos indivíduos por grau de instrução .....	33
Gráfico 7: Gráfico de dispersão da percentagem de indivíduos que compra <i>online</i> em função do grau de instrução.....	34
Gráfico 8: Percentagem de compradores <i>online</i> em cada classe etária.....	35
Gráfico 9: Produtos comprados por mulheres e homens .....	36
Gráfico 10: Representação da variância das dimensões associadas aos produtos comprados <i>online</i> .....	37
Gráfico 11: Coordenadas dos centroides da ACM para os produtos comprados <i>Online</i> : Dimensão 1 vs Dimensão 2 .....	40
Gráfico 12: Coordenadas dos centroides da ACM para os produtos comprados <i>Online</i> : Dimensão 2 vs Dimensão 3 .....	40
Gráfico 13: Coordenadas dos centroides da ACM para os produtos comprados <i>Online</i> : Dimensão 3 vs Dimensão 4 .....	41
Gráfico 14: Coeficientes de fusão da análise de <i>clusters</i> associada aos produtos comprados <i>online</i> – critério vizinho mais afastado.....	42
Gráfico 15: Curva ROC associada ao modelo de regressão logística para o sexo em função dos produtos comprados <i>online</i> .....	49
Gráfico 16: Serviços aderidos pelas mulheres e pelos homens .....	51
Gráfico 17: Representação da variância das dimensões associadas aos serviços aderidos <i>online</i> .....	52
Gráfico 18: Representação do espaço dos serviços aderidos <i>online</i> - Dimensão 1 vs Dimensão 2.....	54
Gráfico 19: Coeficientes de fusão da análise de Clusters associada aos serviços aderidos <i>online</i> .....	55
Gráfico 20: Curva ROC associado ao Modelo de Regressão Logística obtido para os serviços aderidos Online, usando como variável suplementar “Sexo” .....	58
Gráfico 21: Período do dia predilecto para comprar <i>online</i> .....	60
Gráfico 22: Meios de publicidade online que já motivaram compra .....	61
Gráfico 23: Motivos pelos quais os inquiridos não compram <i>online</i> .....	62
Gráfico 24: Distribuição do grau de segurança nas compras <i>online</i> por sexo.....	63
Gráfico 25: Percentis do grau de segurança nas compras <i>online</i> , por idade categorizada .....	64
Gráfico 26: Caixa de bigodes “Grau de Instrução” vs. “Quão seguro considera ser comprar <i>online</i> ?” .....	64
Gráfico 27: Caixa de bigodes do Grau de segurança nas compras <i>online</i> em função da auto-percepção de literacia no uso da <i>internet</i> .....	65
<b>Gráfico 28:</b> Fatores de maior segurança nas compras <i>online</i> .....	65
Gráfico 29: Vantagens associadas às compras <i>online</i> identificadas por homens e mulheres.....	66
Gráfico 30: Vantagens associadas às compras <i>online</i> identificadas por compradores e não compradores ...	67
Gráfico 31: Dendograma das vantagens usando o método dos centroides.....	68
Gráfico 32: Dendograma das vantagens usando o método do vizinho mais afastado .....	68
Gráfico 33: Dendograma das vantagens usando o método de Ward .....	68
Gráfico 34: Representação da variância das dimensões associadas às vantagens da compra <i>online</i> .....	69

Gráfico 35: Coordenadas dos centroides da ACM para as vantagens da compra <i>online</i> : Dimensão 1 vs Dimensão 2 .....	72
Gráfico 36: Coordenadas dos centroides da ACM para as vantagens da compra <i>online</i> : Dimensão 2 vs Dimensão 3 .....	72
Gráfico 37: Coeficientes de fusão da análise de Clusters associada às vantagens da compra <i>online</i> .....	73
Gráfico 38: Curva ROC associado ao Modelo de Regressão Logística obtido para as vantagens na compra Online, usando como variável suplementar “Costuma realizar compras online” .....	79
Gráfico 39: Curva ROC associado ao Modelo de Regressão Logística para o sexo em função das vantagens na compra <i>online</i> .....	81
Gráfico 40: Desvantagens associadas às compras <i>Online</i> identificadas por homens e mulheres.....	83
Gráfico 41: Desvantagens associadas às compras <i>Online</i> identificadas por Compradores e Não Compradores .....	83
Gráfico 42: Dendograma das desvantagens usando o método dos centroides .....	85
Gráfico 43: Dendograma das desvantagens usando o método do vizinho mais afastado.....	85
Gráfico 44: Dendograma das desvantagens usando o método de Ward.....	85
Gráfico 45: Dendograma das desvantagens usando o método de Ward para os inquiridos do sexo Feminino .....	86
Gráfico 46: Dendograma das desvantagens usando o método de Ward para os inquiridos do sexo Masculino .....	86
Gráfico 47: Dendograma das desvantagens usando o método de Ward para os inquiridos não compradores .....	86
Gráfico 48: Dendograma das desvantagens usando o método de Ward para os inquiridos compradores... ..	86
Gráfico 49: Representação da variância das dimensões associadas às desvantagens da compra <i>online</i> .....	87
Gráfico 50: Coordenadas dos centroides da ACM para as desvantagens da compra <i>Online</i> : Dimensão 1 vs Dimensão 2 .....	90
Gráfico 51: Coordenadas dos centroides da ACM para as desvantagens da compra <i>Online</i> : Dimensão 3 vs Dimensão 4 .....	90
Gráfico 52: Coeficientes de fusão da análise de Clusters associada às desvantagens da compra <i>online</i> .....	91
Gráfico 53: Curva ROC associada ao Modelo de Regressão Logística obtido para as desvantagens na compra Online, usando como variável suplementar “Costuma realizar compras online” .....	97
Gráfico 54: Curva ROC associado ao Modelo de Regressão Logística obtido para as desvantagens na compra Online, usando como variável suplementar “Sexo” .....	99
Gráfico 55: Tempo máximo para entrega de compra <i>Online</i> .....	102
Gráfico 56: Inquiridos que comparam produtos <i>Online</i> .....	102
Gráfico 57: Meios de publicidade considerados mais eficazes por cada um dos sexos.....	103
Gráfico 58: Meios de publicidade considerados mais eficazes por Compradores e Não Compradores <i>Online</i> .....	104
Gráfico 59: Dendograma dos meios de publicidade mais eficazes usando o método dos centroides.....	105
Gráfico 60: Dendograma publicidade mais eficazes usando o método da maior distância.....	105
Gráfico 61: Dendograma publicidade mais eficazes usando o método de Ward.....	105
Gráfico 62: Dendograma das desvantagens usando o método de Ward para os inquiridos do sexo Feminino .....	106
Gráfico 63: Dendograma das desvantagens usando o método de Ward para os inquiridos do sexo Masculino .....	106

Gráfico 64: Dendograma das desvantagens usando o método de Ward para os inquiridos não compradores .....	106
Gráfico 65: Dendograma das desvantagens usando o método de Ward para os inquiridos compradores. ....	106
Gráfico 66: : Representação da variância das dimensões associadas aos meios de publicidade mais eficazes .....	107
Gráfico 67: Coordenadas dos centroides da ACM para os meios de publicidade mais eficazes: Dimensão 1 vs Dimensão 2 .....	110
Gráfico 68: Coordenadas dos centroides da ACM para os meios de publicidade mais eficazes: Dimensão 3 vs Dimensão 4 .....	110
Gráfico 69: Coeficientes de fusão da análise de Clusters associada aos meios de publicidade mais eficazes – método do vizinho mais afastado .....	111
Gráfico 70: Curva ROC associado ao Modelo de Regressão Logística obtido para os meios de publicidade mais eficazes, usando como variável suplementar “Costuma realizar compras online” .....	117
Gráfico 71: Curva ROC associado ao Modelo de Regressão Logística para o sexo em função dos meios de publicidade mais eficazes.....	119

# Lista de notações e abreviaturas

$\sigma$  : Desvio-padrão da População

$s$  : Desvio-padrão da amostra

ACM : Análise de Correspondências Múltiplas

AD : Análise Discriminante

RL : Regressão Logística

$\bar{\mu}$  : Média da População

$\bar{x}$  : Média da População

AUC: Área abaixo da curva ROC (Area under the curve)

OR: Odds ratio

# 1) Introdução

Acompanhando o crescimento exponencial dos utilizadores da *Internet* em todo o mundo, também o número de utilizadores do comércio eletrónico ou ***e-commerce*** evoluiu consideravelmente.

O *e-commerce* é uma das **realidades atuais** mais discutidas no **mundo dos negócios** e, enquanto ferramenta de comércio, tem sido responsável pela reformulação das relações entre clientes e fornecedores, pela racionalização dos processos de negócio e pela reestruturação de várias indústrias.

Em **Portugal**, ainda que a proporção de utilizadores que compra *online* seja reduzida, o fenómeno do *e-commerce* tem vindo a crescer na última década e já faz parte do dia-a-dia de muitos Portugueses, complementarmente e/ou em alternativa ao comércio tradicional.

Em termos de investigação o estudo dos modelos de negócio da economia digital é recente e, ainda, existem poucos dados científicos quantificáveis. No entanto, cada vez mais as empresas procuram compreender as **melhores práticas** e observar as **tendências emergentes** que moldam o **comportamento do consumidor online**.

Tendo em consideração o papel central que o *e-commerce* já desempenha no comércio português e o seu potencial de influência das práticas de consumo, este trabalho procura, em termos gerais, analisar o comércio eletrónico estabelecido entre empresas e consumidores finais individuais portugueses.

Começamos por enquadrar o *e-commerce* na evolução da *internet* e por analisar a sua própria evolução, assim como os modelos propostos para o comportamento do consumidor *online*. De seguida, com base na literatura disponível, traçamos o perfil sociodemográfico do consumidor *online*, as motivações e vantagens que associa ao *e-commerce*, assim como as perceções de risco, nomeadamente no que diz respeito à segurança e privacidade. Efetuamos ainda uma síntese de alguns estudos relevantes já realizados em Portugal sobre este tema. No final deste capítulo introdutório, justificamos a pertinência do estudo que foi realizado bem como dos seus objetivos. No capítulo seguinte fazemos uma breve revisão de alguns métodos estatísticos que serão utilizados na análise do conjunto de dados recolhidos. Posteriormente, descrevemos a metodologia utilizada para definir a amostra, a elaboração do questionário e como procedemos à recolha dos dados. Finalmente, apresentamos e discutimos os resultados obtidos, terminando com a enunciação de algumas considerações finais.

## 1.1) Evolução da *Internet*

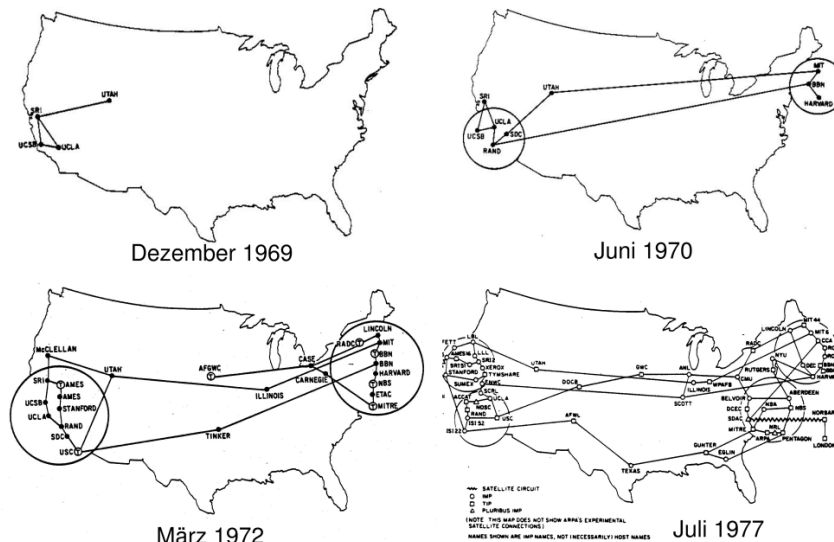
A sociedade atual vive num mundo onde a troca e a partilha de informação de conteúdos são uma constante do dia-a-dia. Em 2015, utilizando apenas um *smartphone* com acesso à *internet*, é possível trocar *emails*, enviar fotografias ou vídeos, pesquisar e discutir sobre temas de interesse, ver e falar com pessoas

que estão no outro lado do mundo, conhecer novas pessoas, reservar voos e hotéis, acompanhar o que estão os amigos a fazer nesse momento ou mesmo gerir as finanças pessoais sem grande dificuldade.

O caminho para chegar aqui não foi fácil, iniciou-se com o clima de tensão gerado após o final da Segunda Guerra Mundial e que deu origem à Guerra Fria – período no qual eram comuns as ações de espionagem entre os Estados Unidos e a União Soviética. Como reação ao lançamento efetuado pela Rússia do Sputnik - o primeiro satélite artificial lançado para o espaço exterior – os Estados Unidos criaram, em 1957, a ARPA (Advanced Research Project Agency) com o objetivo de desenvolver programas relacionados com satélites e o espaço. Um ano mais tarde com a criação da NASA (National Aeronautics and Space Administration), a ARPA perdeu espaço de atuação e centrou-se na investigação em Informática.

Em Outubro de 1958, Joseph Licklider, contratado pela ARPA, criou o Information Processing Techniques Office (IPTO), de modo a garantir a comunicação entre os investigadores, conduziu a investigação para a criação de uma rede – NET – de transferência de dados entre computadores. Licklider abandonou a ARPA em 1965, porém a investigação nesta área continuou a ser desenvolvida pela ARPA. Após alguns anos com desenvolvimentos na forma de partilha da informação (que levaram a criação de protocolos de transferência de dados) foi criada, em 1969, a primeira rede de computadores – ARPANET – que ligava computadores entre a Universidade da Califórnia em Los Angeles, o Stanford Research Institute, a Universidade de Utah e a Universidade da Califórnia em Santa Bárbara, utilizando a rede telefónica para partilhar a informação, estando a expansão das primeiras *nets* representadas na Figura 1.

**Figura 1:** Mapa com evolução das primeiras redes (*nets*)



Fonte: <https://sites.google.com/site/sitesrecord/o-que-e-arpamet>

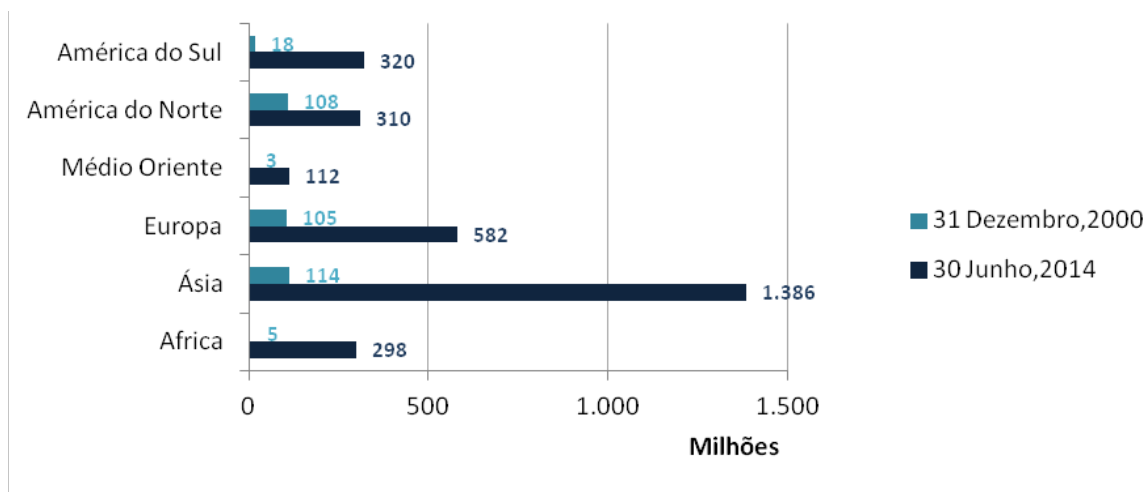


Com o aparecimento de outras redes nos Estados Unidos, foi necessário garantir a comunicação entre várias NETs de informação. Vinton Cerf e Robert Kahn utilizaram o protocolo TCP/IP para garantir a comunicação entre 3 redes distintas (ARPANET, RPNET e STATNET) e aí nasceu o conceito de *internet*. Contudo, este termo só se tornou popular quando, em 1990, a ARPANET foi substituída pela NSFNET, que se popularizou como *internet*. Em 1991 o termo *World Wide Web* é utilizado pela primeira vez.

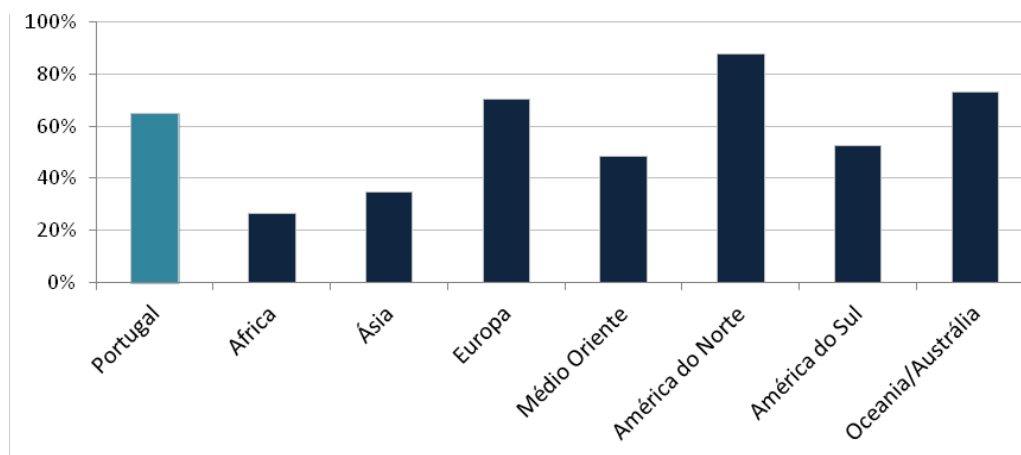
Rapidamente, o surgimento de novas redes estendeu-se ao continente Europeu, com redes em Inglaterra e em França onde no âmbito do projeto Cyclades, foi construída a TRANSPAC – um nó desta rede foi instalado em Lisboa, nascendo assim a *internet* em Portugal. Na década de 80, a *internet* começa a ser utilizada por algumas empresas e universidades portuguesas, sendo que em 1991 todas as universidades estavam abrangidas pela RCCN – Rede da Comunidade Científica Nacional. Em 1995, a Portugal Telecom surge como o primeiro *Internet Service Provider* (ISP) e lança o *kit* Telepac, democratizando assim o acesso à *internet* à população em geral. Desde então surgiram vários novos ISP's, e a tecnologia sobre o qual os dados são transferidos evoluiu desde ligações por *modem* e RDIS (Rede Digital Integrada de Serviços vulgo linha telefónica), a ligações por ADSL, cabo ou mais recentemente fibra ótica, nas quais a velocidade de transferência de dados é bastante superior – atualmente existem soluções com 400 Mbps de velocidade garantida de transferência, que comparam com 56 Kbps de velocidade máxima dos modems da década de 90.

Na primeira década do século XXI, com a democratização da *internet* como bem de consumo pessoal e doméstico, assistiu-se àquilo que é vulgarmente chamada “explosão da *internet*”, aumentando exponencialmente o número de utilizadores, de *sites*, de novos conteúdos e de novas aplicações da *internet* (Gráficos 1 e 2). Foi, a partir deste período, que surgiram alguns dos principais serviços que se tornaram referência e parte integrante do dia-a-dia de milhões de pessoas da atualidade (Tabela 1).

**Gráfico 1:** Utilizadores de *internet* por continente



Fonte: <http://www.internetworldstats.com>

**Gráfico 2:** Penetração de Internet na População (Junho/2014)

Fonte: <http://www.internetworldstats.com>

**Tabela 1:** Principais lançamentos da história da *internet* nos primeiros anos do século XXI

Ano de lançamento	Serviço	Descrição
2001	Wikipedia	Repositório de conhecimento
2003	Wordpress	Ferramenta de criação e edição de <i>sites</i>
2003	iTunes Store	Loja <i>online</i> de música e filmes
2003	LinkedIn	Rede social para profissionais
2003	Skype	Ferramenta para videoconferências
2004	Facebook	Rede social
2004	Gmail	Ferramenta de <i>email</i>
2005	Youtube	Ferramenta de partilha e visualização de vídeos
2006	Twitter	Rede social
2008	Evernote	Ferramenta de organização de informação
2010	Pinterest	Rede social
2012	Google +	Rede social

Atualmente, em 2015, considera-se que estamos a viver a 2ª geração da *internet* (ou Web 2.0). A primeira geração da *internet* (ou Web 1.0) diz respeito à implementação da rede *per se* e do desenvolvimento de *websites*. A segunda geração é caracterizada pelo relacionamento dos utilizadores entre si, gestores de conteúdos ou empresas – tendo como exemplo a relação dos utilizadores com o Youtube ou o Facebook ou ferramentas colaborativas como a Wikipedia, nos quais são os próprios utilizadores a gerarem os conteúdos do próprio *website*. A terceira (e próxima) geração de *internet* (Web

3.0 ou Web semântica) tem por base a premissa de que a rede se consiga organizar autonomamente, “aprender sozinha” com o conhecimento disponível e apresentar as melhores soluções para um utilizador específico sobre determinado problema. O conceito associado a esta 3ª geração de *internet* tem sido bastante criticado e posto em causa por ser um conceito extremamente ambicioso, idealista e que dificilmente será executado. Ainda assim, os gigantes da indústria já estão a preparar os seus motores de busca para esta terceira geração, sendo que o Google Search, o Microsoft Bing e o Wolfram Alpha, de Stephen Wolfram têm dado passos importantes nesta matéria.

Outro avanço que importa destacar relativamente ao modo como é consumida a *internet* é a multiplicidade de dispositivos e dos novos usos que é possível dar à *internet*. Longe vão os anos em que era necessário um computador *desktop* ou um portátil para aceder à *internet*. Neste momento, e a partir de preços bastante reduzidos e ao alcance da generalidade da população, é possível a aceder à *internet* através de *smartphones*, *tablets* ou *wearable devices*. Estes últimos dispositivos, que estão ligados à *internet* através de rede *wireless* ou móvel e são, muitas vezes, usados como um acessório de moda (pulseiras, relógios, fitas de competição), constituem uma tendência crescente nos últimos anos, permitindo-nos monitorizar vários aspetos da vida quotidiana (por exemplo, o número de kms percorridos durante uma corrida matinal, o número e a qualidade das horas de sono, etc.) e prometem revolucionar a forma como experienciamos o dia-a-dia.

## 1.2) Evolução do *E-commerce*

Em termos gerais, o termo *e-commerce* (*Electronic Commerce*), que em português significa Comércio Eletrónico, corresponde às transações comerciais efetuadas através de dispositivos eletrónicos com acesso à *internet* (computadores, *smartphones*, *tablets*, ...), a partir das quais são adquiridos produtos ou serviços através de meios puramente digitais, que não envolvem contacto humano entre o consumidor e o comerciante.

Turban *et al.* (2008) definem o comércio eletrónico como o processo de compra, venda, transferência ou troca de produtos, serviços, e/ou informação através de redes de computadores, incluindo a *internet*. Já o Eurostat adota uma definição em linha com a defendida pela Organização para a Cooperação e Desenvolvimento Económico (OCDE), segundo a qual o comércio eletrónico consiste na transação de bens e serviços entre computadores mediados por redes informáticas, com a ressalva de que o pagamento ou a entrega dos produtos transacionados não têm de ocorrer, obrigatoriamente, através dessas redes.

Independentemente da forma como o conceito de *e-commerce* é definido, importa reter que aquilo que distingue o comércio eletrónico do comércio tradicional é, sobretudo, a forma como a informação é trocada e processada entre as partes intervenientes, sendo que, no caso do comércio eletrónico, o contacto pessoal direto entre os envolvidos é substituído por fluxos de informação transmitidos através de

uma rede digital ou de outro canal eletrónico (ANACOM, 2004).

Existem diversas formas de classificar o *e-commerce*, consoante, por exemplo, o tipo de produto ou serviço transacionado, o sector de atividade a que correspondem, a tecnologia de suporte usada, os montantes envolvidos nas transações ou o tipo de intervenientes no processo. Uma das classificações mais comuns tem em conta o tipo de intervenientes envolvidos e identifica quatro tipos principais de comércio eletrónico (Saraiva, 2012): *Business-to-Business* (B2B), *Business-to-Consumer* (B2C), *Business-to-Administration* (B2A) e, por último, *Consumer-to-Administration* (C2A). Neste trabalho analisaremos o tipo de *e-commerce* B2C, uma vez que pretendemos estudar os hábitos de *e-commerce* dos consumidores portugueses.

Antes da existência destas classificações, a história do *e-commerce* nasce com Michael Aldrich, que em 1979 inventou o conceito de *Online Shopping* para os segmentos B2C e B2B. Em 1980, Aldrich inventa ainda o Teleputer (muitas vezes traduzido como “Telecomputador”) – uma máquina que resulta da fusão de um televisor, um computador e a própria rede de telefone doméstica. O Teleputer foi produzido e comercializado durante vários anos como uma máquina para fazer transações eletrónicas. Em 1981 a Thomson Holidays faz a primeira transação eletrónica da história, neste caso no mercado B2B. Três anos mais tarde, a Tesco faz a primeira venda para o segmento B2C.

Em 1994 foram feitos vários avanços ao nível da encriptação de dados (SSL Encryption), o que levou a uma maior segurança nos pagamentos feitos através da *internet*. Nesse mesmo ano, a PizzaHut vendeu a sua primeira *pizza online*. Em 1995, nasceram dois dos principais *players* do mercado de *e-commerce* da atualidade: a Amazon e o eBay.

Em 2002 o eBay adquire o PayPal, ferramenta que foi adotada pela generalidade dos *sites* de *e-commerce*, e que tornou as compras *online* mais rápidas, cómodas e seguras. Mais recentemente é possível a qualquer utilizador que tenha um *site* e um negócio próprio ter a sua própria loja *online* quer através de soluções *open-source*, quer através de ferramentas como o Shopify que faculta, através do pagamento de uma mensalidade, uma loja *online* com as funcionalidades mais avançadas do mercado, não exigindo conhecimentos técnicos de informática avançados por parte do utilizador para colocar o seu negócio *online*. Outra das soluções encontradas por pequenos retalhistas foi a venda dos seus produtos/serviços através de ferramentas colaborativas ou redes sociais, como o Facebook, através das quais os comerciantes expõem gratuitamente os seus produtos/serviços para uma determinada audiência.

Desde o início da primeira década de 2000 até hoje o comércio eletrónico B2C tem registado uma evolução assinalável, quer no espaço europeu, quer em Portugal, demonstrando o enorme potencial que o comércio *online* tem na sociedade dos nossos dias. Entre 2004 e 2010, a proporção de consumidores que realizou compras *online* nos últimos 12 meses cresceu significativamente, duplicando (aumento de 20% para 40%) no caso da União Europeia e triplicando (aumento de 5% para 15%) no caso de Portugal (Santos, 2011). Ainda assim, na generalidade dos países da União Europeia a proporção de consumidores que efetua

compras *online* é relativamente baixa sendo, na generalidade dos países, inferior a 50%. Da mesma forma, a frequência com que a maior parte dos consumidores realiza compras *online* é reduzida (Santos, 2011).

Já no que diz respeito à percentagem das empresas que vende *online*, de acordo com um relatório da Comissão Europeia (2009), Portugal está acima da média da União Europeia. Apesar do *e-commerce* ainda estar numa fase de crescimento, as empresas já implementam este sistema de comércio e compreendem as suas mais-valias (Bhanji, 2010).

Por último, importa destacar alguns fatores que têm tido um papel de relevo no crescimento significativo das vendas em *e-commerce*, ao contrário do que acontece com o comércio tradicional:

- O aumento da velocidade de acesso à *internet*, a hiper conectividade que o serviço móvel possibilita e a democratização da tecnologia;
- As vendas através de *sites* de promoções ou vales promocionais;
- A evolução dos métodos de pagamento;
- O lançamento, em 2000, do Google AdWords – uma ferramenta integrada na pesquisa do Google que apresenta anúncios associados à pesquisa do utilizador. Esta ferramenta permitiu às empresas apresentar os seus produtos numa zona de destaque no momento em que o utilizador está a fazer pesquisas relacionadas com esses produtos, pagando apenas caso o utilizador clique no anúncio.

### 1.3) Modelos de Comportamento do Consumidor *Online*

Vários autores têm procurado adaptar e criar modelos de comportamento do consumidor adequados ao contexto *online*, que podem ajudar a caracterizar e a compreender os hábitos de *e-commerce*. Cheung *et al.* (2003) propuseram um Modelo de Intenção, Adoção e Continuidade/Recompra. De acordo com este modelo existem 5 fatores que moldam a intenção, a adoção e a continuidade da compra *online*:

- *Características do consumidor*: as características demográficas, de personalidade, estilo de vida, motivação, satisfação, confiança, entre outras determinam e revelam muito sobre a adesão à compra *online*;
- *Características do produto/serviço*: o tipo de produto, a frequência de compra, a tangibilidade e a diferenciação de preço são algumas das características que moldam a perceção do consumidor *online*;
- *Características do meio*: o consumidor *online* valoriza e procura a facilidade de uso do meio, a segurança e a rapidez;
- *Características do comerciante e do intermediário*: na escolha do consumidor *online* também pesam fatores como a qualidade do serviço, a privacidade dos dados, o sistema de entrega, o serviço pós-venda ou a reputação da marca;

- *Influências do ambiente envolvente*: a cultura, a imagem, o ambiente social ou os hábitos de trabalho são outros dos elementos que podem influenciar o consumidor *online*.

Zhou *et al.* (2007) propõem outro modelo – o Modelo de Aceitação de Compras *Online*. Os autores analisaram a literatura existente sobre os fatores de consumo que influenciam o comportamento do consumo *online*. De acordo com este modelo a experiência na *internet* afeta as motivações e a intenção de compra *online*, assim como a experiência do consumidor online. As características demográficas dos consumidores (como a cultura, o sexo, a idade ou tipo de produtos que compram) têm impacto sobre a percepção de risco e o resultado percebido da compra – elemento que assume importância neste modelo. Este resultado refere-se à percepção de possíveis resultados (positivos ou negativos) do comportamento de compra. A intenção e a realização da compra *online* são ainda afetadas pelas crenças normativas dos consumidores e pelas recomendações de família e amigos.

A satisfação, enquanto sentimento geral do consumidor *online* sobre a experiência anterior de compras *online*, constitui um forte indicador da continuidade da intenção e realização de compras *online*, a longo prazo (Saraiva, 2012).

## 1.4) Perfil Sociodemográfico do Consumidor *Online*

De acordo com os resultados de 2009 do Bareme Internet da Marktest, um em cada cinco portugueses já fez compras de produtos ou serviços através da *internet*. Estes 1 667 mil portugueses representam 20,1% do total de indivíduos, com idade superior a 15 anos, que residem no continente, e 37,2% dos utilizadores da *internet* (Bhanji, 2010). Apesar de 83% dos utilizadores pesquisar informação sobre bens e serviços na *internet*, apenas 34% destes concretiza a realização de compras *online* (Santos, 2011). Que características diferenciam estes consumidores *online*?

De acordo com Santos (2011) os indivíduos do sexo masculino, com idades compreendidas entre os 16 e os 44 anos, detentores de habilitações literárias ao nível do ensino secundário e do superior, estudantes e empregados, assim como residentes na região de Lisboa destacam-se positivamente da proporção média nacional de utilizadores do comércio eletrónico. Segundo Bhanji (2010) os idosos, residentes no Interior Norte, reformados e domésticas, assim como consumidores de classes sociais mais baixas são os que menos aderem ao *e-commerce*.

Saraiva (2012) sintetizou os principais resultados de diversos estudos que procuraram investigar os fatores sociodemográficos gerais que influenciam as compras *online*, concluindo que os homens fazem compras *online* com mais frequência e gastam mais dinheiro do que as mulheres; não existem resultados claros sobre a relação entre a idade, o grau de instrução e frequência de compra *online*.

No que diz respeito ao género também Rodgers e Harris (2003 cit. *in* Ferreira, 2013) afirmam que apesar

do número de utilizadores da *internet* estar igualmente dividido entre géneros, mais homens do que mulheres realizam compras *online*. Em Portugal, segundo dados do INE/UMIC, em 2012, a utilização do *e-commerce* era muito semelhante entre homens e mulheres (14,1% e 12,5%, respetivamente).

Estudos recentes sublinham que homens e mulheres valorizam aspetos diferentes do *e-commerce*: as mulheres focam-se mais na confiança e garantia de segurança, assim como na capacidade de partilhar opiniões; os homens interessam-se mais pelo valor obtido através da compra (Ferreira, 2013). Para os homens não é relevante o facto de no consumo *online* não ser possível interagir fisicamente com o vendedor, mas privilegiam a conveniência das compras *online*. Para as mulheres, acontece o oposto: dão mais importância à dimensão social da compra do que à conveniência. Este pode ser um dos fatores que explica o facto e utilizarem o *e-commerce* com menor frequência (Saraiva, 2012).

Relativamente à idade os estudos são contraditórios em relação à idade dos consumidores *online* no que diz respeito à intenção e frequência da compra *online* – alguns estudos identificam relações positivas, outros negativas e outros ainda uma ausência de relação (Saraiva, 2012).

Um estudo realizado pela ACEPI (2010) notou que a categoria de produtos mais comprados através da *internet* durante o primeiro trimestre de 2010 foi o da Eletrónica/Telemóveis (24%), seguido pela Informática (15%) e Alimentação/Bebidas/Tabacaria/Charutos (15%).

## 1.5) Motivações e Vantagens Associadas ao *E-commerce*

Uma vez que, no contexto *online*, o consumidor interage com elementos diferentes do contexto offline, nomeadamente com a tecnologia, é importante compreender as preferências e motivações do consumidor para realizar compras através do comércio eletrónico.

Analisando diversos estudos, Lima (2012) concluiu que o comportamento de compra *online* é motivado por vários fatores, entre os quais se destacam o preço baixo, a confiança, a conveniência, a poupança de tempo, o atendimento ao consumidor e a possibilidade de comparação de produtos.

Ferreira (2013) reúne um conjunto de investigações que distinguem entre motivações utilitárias (consumidor orientado por objetivos – quer cumprir o objetivo da compra de modo eficiente e sem desperdiçar tempo) e hedónicas (consumidor orientado por experiências – procura sensações de prazer, diversão e fantasia durante a compra) para a compra *online*.

A revisão de literatura realizada por Saraiva (2012) confirma que os consumidores compram *online* pelo seu carácter utilitário e vêm na conveniência o principal benefício associado à compra *online*, nomeadamente ao nível do acesso, poupança de tempo e eficiência.

Castro (2012) estudou, em Portugal, os fatores que influenciavam a procura de informação e a intenção de compra *online* de seguro automóvel. Na procura de informação obteve evidências de que as variáveis com maior impacto são a perceção de utilidade, a atratividade de preço, o envolvimento e a perceção de

facilidade de uso. Para a intenção de compra *online*, as variáveis que revelaram ter maior impacto foram a percepção de utilidade, a atratividade de preço, a confiança e a lealdade.

É de salientar que, tal como no mercado *offline*, também no mercado *online* a credibilidade da empresa é crucial – os consumidores *online* compram as marcas que conhecem e nas quais confiam (Maya e Otero, 2002 cit. in Saraiva, 2012). Da mesma forma a personalização é muito importante na qualidade da experiência de compra *online* e, portanto, as campanhas de comunicação, a interação comercial e as situações de reclamação devem ser sempre geridas de forma orientada para o relacionamento com o cliente (Saraiva, 2012).

Tendo em conta os fatores que parecem explicar a motivação para o consumo *online*, vários estudos e o contributo de diferentes autores, Santos (2011) enumera os seguintes benefícios e vantagens do *e-commerce*:

- *Conveniência*: possibilidade de comprar produtos e serviços em qualquer momento e em qualquer local; acesso a um mercado global, sem barreiras geográficas;
- *Redução de custos*: ao permitir reduzir os custos de comunicação, recolha de informação, processamento de transações, distribuição de produtos digitais, cadeia de abastecimento e serviço ao cliente, a *internet* possibilita a prática de preços mais baixos no comércio eletrónico;
- *Qualidade do serviço*: os consumidores *online* beneficiam de uma melhor qualidade do serviço e de um apoio pré e pós-venda mais eficiente, dada a proximidade entre fornecedores e clientes;
- *Personalização e Foco no consumidor*: o comércio eletrónico permite aos fornecedores recolher informações relevantes sobre as características, gostos, necessidades, desejos e comportamentos do consumidor *online*. O uso dessa informação permite conceber uma oferta de bens e serviços adequada ao perfil do cliente e utilizar estratégias de *marketing one-to-one* que conduzem a uma verdadeira personalização;
- *Ausência de pressão para comprar*: uma vez que não existem vendedores persuasivos, o consumidor pode tomar decisões ao seu ritmo e sem pressão.

Resta acrescentar que o mercado *online* tem também um papel importante na tomada de decisões de consumidores *offline*. Muitos consumidores, embora finalizem o processo de compra *offline*, recorrem a *sites online* para comprar preços e procurar informações sobre os produtos (Broekhuizen e Huizingh, 2009 cit. in Bhanji, 2010).



## 1.6) Percepções de Risco Associadas ao *E-commerce*

A percepção de risco consiste na crença subjetiva do consumidor poder ter uma má experiência de compra, ou seja, a compra poder não corresponder ao que é esperado. Esta percepção de risco influencia negativamente a intenção de compra e a realização efetiva de compras pelo consumidor *online* e é maior do que a percepção de risco associada ao comércio tradicional, nomeadamente porque não existe uma interação direta e humana na relação comercial, assim como pelos riscos de privacidade e partilha de dados que a utilização da *internet* acarreta. Lima (2012) concluiu que os consumidores *online* fariam mais compras online não fossem as preocupações ao nível dos riscos de quebra de segurança e violação da privacidade. Bhatnagar *et al.* (2000 cit. in Saraiva, 2012) vão mais longe e referem que a percepção de risco nas compras *online* é maior no caso de produtos do que no caso de serviços.

Santos (2011) sublinha que, em Portugal, a preocupação do consumidor com a privacidade da informação pessoal surge como um dos principais motivos para a não adoção do comércio eletrónico. A não realização de compras *online* é referida pelos consumidores como uma das consequências mais relevantes da preocupação com a segurança. De acordo com o autor, em Portugal, 57% dos utilizadores da *internet* que não compraram *online* no último ano fizeram-no por preocupações com a segurança, 53% por preocupações com a privacidade e 42% por preocupações com a confiança.

Desta forma, a segurança e a privacidade constituem desafios essenciais para o sucesso do *e-commerce*. Lopes (2014) analisou vários estudos sobre os fatores que mais influenciam a percepção de segurança do consumidor *online* e enumerou os seguintes: atitude para com a segurança e propensão do consumidor para confiar; conhecimento e experiência dos recursos de segurança; facilidade de uso do interface e apresentação do *website*; presença de “https” na barra de endereço; existência de uma política de privacidade e segurança e de receção eletrónica; credibilidade e certificação por terceiros da empresa que fornece o serviço ou produto.

Bhanji (2010) sublinha ainda que aquando do momento de realização de uma compra *online*, o consumidor se questiona sobre a forma de pagamento e os mecanismos de armazenamento e transmissão de informações pessoais. O consumidor pode evitar a compra *online* devido à percepção dos riscos envolvidos na transmissão de informações como o número de cartão de crédito.

Lee e Turban (2001 cit. in Santos, 2011) conceberam um modelo teórico que inclui quatro grandes determinantes que estabelecem a confiança *online* e são moderados pela propensão individual para confiar:

- Confiança no retalhista *online* (capacidade, integridade e benevolência);
- Confiança na *internet* como meio de compra;
- Infraestruturas (segurança e certificação por entidades externas, por exemplo);
- Outros fatores (como a dimensão da empresa ou variáveis socio demográficas).

## 1.7) Meios de Publicidade *Online*

A publicidade *online* corresponde às ações de *marketing* através de meios interativos, nomeadamente a *internet*, e utiliza um tipo de comunicação particularmente orientado para a interação (Nordeste, 2009).

De acordo com Zeff e Aronson (2000 cit. in Cabral, 2011), a publicidade *online* possui quatro características distintas: focalização (possibilidade de impactar determinados consumidores ou áreas geográficas, assim como direcionar as mensagens promocionais conforme a hora do dia ou a plataforma *online*); monitorização (do interesse e do envolvimento do utilizador com determinada marca, assim como da resposta a determinada peça publicitária através do número de vezes que é clicada, do número de compras ou de indicações que gerou); entrega e flexibilidade da mensagem (possibilidade de anunciar em tempo real, 24 horas por dia, 365 dias por ano, assim como de atualizar ou cancelar a campanha a qualquer momento); e interatividade (possibilidade dos consumidores interagirem e comprarem um produto).

À medida que a *internet* se desenvolve vão surgindo novas formas e formatos de publicitar *online* – *banners*, botões, *floating ads*, *interstitial banners*, *map ads*, MRECS, *pop-up ads* ou *superstitials*. Para além destes, mais tradicionais, têm surgido abordagens de *marketing one-to-one*, que personalizam as mensagens publicitárias conforme os clientes – *email marketing*, *newsletters*, *emails* promocionais, *marketing viral* e *Search Engine Marketing*, por exemplo (Nordeste, 2009).

As campanhas de *email marketing* constituem um dos meios de angariação comercial mais utilizados atualmente, permitindo uma abordagem direta, sem custos (ou com custos muito reduzidos) que atinge milhares de utilizadores recorrendo a uma taxa de esforço diminuta. Face a outros canais de publicidade *offline* (como a televisão ou os *outdoors*), o *email marketing* permite às empresas medir o impacto efetivo da Ação publicitária (por exemplo, número de utilizadores que viram o email, que o abriram, que foram ao *site* da empresa ou que acabaram por realizar uma compra na sequência do *email*).

É possível ainda, hoje em dia, associar publicidade aos motores de busca ou às redes sociais. Nestes casos, as empresas podem publicitar os seus produtos ou serviços apenas a determinada audiência e ainda recolher informação detalhada sobre os utilizadores de modo a, posteriormente, adequar novas ações publicitárias aos seus gostos, interesses e comportamentos de consumo (Nordeste, 2009).

Por exemplo, um anunciante que lance uma ação publicitária no Facebook pode segmentá-la de modo muito específico em função do perfil de consumidor que pretende atingir: mulheres, entre os 20 e os 60 anos, que residam no centro do país, que já sejam clientes da marca. Ou ainda em função do seu histórico de navegação na *internet*, por isso, é frequente, quando pesquisamos um determinado hotel, voltarmos a ver anúncios a esse mesmo hotel na página de Facebook ou noutras que permitam este tipo de ação publicitária de *remarketing*.

## 1.8) Estudos Realizados sobre *E-commerce* em Portugal

Santos (2011) procurou analisar a forma como a preocupação com a privacidade da informação pessoal na *internet* exerce o seu impacto na intenção do consumidor realizar compras *online*. Construiu um modelo de investigação que estabelece que a preocupação com a privacidade exerce o seu efeito na intenção de comprar *online*, por intermédio da sua ligação ao modelo da confiança-risco, à teoria do comportamento planeado e ao modelo de adoção de tecnologia. Os principais resultados evidenciam que a preocupação com a privacidade na *internet* exerce o seu efeito em diversas crenças sobre a utilização do comércio eletrónico. Este efeito é negativo sobre a confiança, a utilidade percebida, a facilidade de utilização percebida e o controlo do comportamento percebido, sendo positivo sobre o risco percebido. Por sua vez, as três primeiras crenças, de uma forma positiva, e o risco percebido, de uma forma negativa, têm uma influência sobre a atitude em relação à utilização do comércio eletrónico. A intenção de utilizar o comércio eletrónico, enquanto principal variável explicada do modelo, sofre o impacto positivo da utilidade percebida, do controlo do comportamento percebido e da atitude (Santos, 2011).

Com base numa amostra de 343 consumidores, obtida por uma amostragem do tipo bola de neve (através de *email* e Facebook), Saraiva (2012) procurou identificar alguns dos principais preditores que levam os consumidores a realizar compras *online*, nomeadamente as características demográficas, as motivações e experiências de compra *online*, as crenças normativas e as perceções de risco e de benefícios associados à compra através da *internet*. O autor concluiu que existe influência entre o género e a frequência e montante gasto *online*; que os amigos e os familiares são as referências que mais influência exercem nas compras *online* e que a satisfação *online* influencia positivamente as compras *online*. Registou ainda diferenças nos produtos comprados *online* entre cada um dos géneros, bem como na perceção de risco associados à *internet*.

Ferreira (2013) analisou a existência de divergências entre as perceções dos sexos masculino e feminino relativamente a compras *online*. Através de um questionário *online* inquiriu 274 respondentes, com base numa amostragem do tipo bola de neve (recorrendo ao Facebook para divulgar o questionário), sendo a maioria do sexo feminino (58,8%), entre os 25 e os 34 anos de idade (43,8%), com habilitações literárias equivalentes ou superiores à licenciatura. A dimensão onde foram encontradas maiores diferenças percentuais entre ambos foi a dimensão Emoção (grau emocional de homens e mulheres no momento de compra *online* e grau e satisfação emocional após a compra efetuada), enquanto a dimensão Conveniência (grau de utilidade prática/conveniência atribuído a compras *online*) mostrou ser a que tanto as Mulheres como os Homens inquiridos atribuem maior importância.

Lopes (2014) realizou uma investigação sobre as preferências dos consumidores em mercados digitais e tradicionais. Dos 172 participantes que responderam ao questionário *online*, enviado por *email* para a base da Netsonda, a maioria tinha entre 25 e 34 anos (44%), formação de grau superior (82%) e eram grandes

utilizadores da *internet* (59,3%). Apenas 19% pesquisava produtos e serviços na *internet* e só 11% comprava *online* serviços e produtos. Os participantes compravam mais frequentemente viagens/estadias (52,7%), no Mercado *online* (comparativamente ao mercado tradicional). A maioria dos inquiridos revelou ter um interesse moderado relativamente às campanhas de *marketing*. Acerca dos aspetos mais valorizados na compra de produtos/serviços via *internet* a maior parte dos participantes referiu a possibilidade de pesquisa e comparação de produtos (24%) e a comodidade (22%), assim como preços mais baratos (19%) e rapidez (17%). Sobre os aspetos menos valorizados na compra de produtos/serviços *online*, os participantes referiram o não poder experimentar o produto (16%), ter de pagar portes de envio (15%), a dificuldade em devolver o produto (12%), o prazo de entrega (9%), a insegurança ao fazer o pagamento e na receção da encomenda (8%), o esgotamento do *stock* (7%), a confiança no produto e o onde comprar (6% cada), o não ter diversificação de modos de pagamento (5%). Em resposta à pergunta sobre a frequência de compra *online*, a maior parte dos inquiridos respondeu uma vez por mês (53%) e uma vez por ano (30%). A maior parte dos participantes classifica o seu nível de confiança relativamente às compras *online* como médio-alto.

## 1.9) Pertinência e Objetivos do Estudo

A presente dissertação tem como objetivo geral contribuir para o conhecimento sobre os hábitos de *e-commerce* em Portugal, sob uma perspetiva *Business-to-Consumer*.

Como vimos, a generalização do acesso à *internet* em Portugal, as alterações dos hábitos dos consumidores, bem como a crescente penetração das tecnologias móveis (nomeadamente *smartphones* e *tablets*) conduziram a uma transformação dos processos de compra, posicionando o *e-commerce* como um elemento cada vez mais indispensável no mundo empresarial.

Neste sentido, e na medida em que os resultados permitam conhecer melhor o perfil do consumidor *online*, este estudo pode dar um contributo empresarial relevante.

O conhecimento das características dos hábitos de *e-commerce* pode aumentar a eficácia das estratégias de *e-commerce*, abrindo portas a uma abordagem de segmentação, da qual podem decorrer ações concretas para diferentes perfis de clientes. Conhecer os fatores que determinam (ou não) a realização de compras *online*, pode ainda contribuir para que as empresas possam influenciar mudanças de comportamento nos seus clientes, aumentando a eficiência do plano de comunicação e otimizando processos e estratégias de canal que valorizem o que é mais potenciador do comportamento desejado.

Desta forma, quer o crescimento do *e-commerce* em Portugal, quer a sua relevância para a definição futuras de estratégias de *e-commerce* por parte das empresas, justificam a pertinência do presente estudo.

Mais concretamente, os objetivos específicos ao nível das aplicações aos quais se pretende dar resposta através deste trabalho são:

- Traçar o perfil dos consumidores *online* do ponto de vista:
  - Das suas características sociodemográficas (idade, sexo, distrito de residência, estado civil, número de filhos e grau de instrução);
  - Da sua experiência, tipo de compras e hábitos de *e-commerce*;
  - Das desvantagens e percepções de risco associadas à compra *online*;
  - Das motivações e vantagens associadas à compra *online*.
- Obter informação sobre o comportamento do consumidor online que permita às empresas melhorar a eficácia das suas estratégias de *e-commerce*.

Além disso, do ponto de vista estatístico, pretendem-se comparar os resultados obtidos com a utilização de três técnicas diferentes – a Análise de Correspondências Múltiplas, a Análise Discriminante com particular incidência na função de Fisher e a Regressão Logística – na identificação das características que permitam distinguir os compradores *online* dos não compradores *online*.

## 2) Métodos Estatísticos

Ao longo deste trabalho será privilegiado o uso das técnicas de análise exploratória dos dados, uma vez que habitualmente neste tipo de estudos se usam amostras recolhidas com processos de amostragem não probabilísticos (ver Capítulo 3). Para além da usual estatística descritiva univariada, serão usados coeficientes de correlação para o estudo da eventual relação entre duas variáveis, e na análise multivariada serão usadas as técnicas: Análise de Correspondências Múltiplas, Análise Discriminante (com particular interesse na função de Fisher) e Regressão Logística.

### 2.1) Coeficientes de Correlação

Nos questionários é usual predominarem questões cuja resposta é de natureza qualitativa. É nesse sentido que se equadram os coeficientes de correlação: Phi, de Contingência e de Spearman, que permitem estudar a relação entre duas variáveis qualitativas.

#### Coeficiente de correlação de Phi

O coeficiente de correlação Phi é indicado quando as variáveis são dicotómicas. O coeficiente de correlação Phi obtém-se a partir do coeficiente linear de Pearson, assumindo que as variáveis são do tipo dicotómico. Deste modo, o coeficiente Phi está relacionado com o coeficiente de qui-quadrado para uma tabela de contingência 2x2.

**Tabela 2:** Tabela de contingência

	y=0	y=1	Total
x=0	$n_{0\ 0}$	$n_{0\ 1}$	$n_{0,k}$
x=1	$n_{1\ 0}$	$n_{1\ 1}$	$n_{1,k}$
Total	$n_{k,0}$	$n_{k,1}$	$n$

Considerando uma tabela de contingência do tipo da apresentada na Tabela 2, o coeficiente de correlação de Phi é dada pela seguinte expressão (Maroco,2007):

$$\phi = \frac{n_{00} \cdot n_{11} - n_{01} \cdot n_{10}}{\sqrt{n_{k,0} \cdot n_{k,1} \cdot n_{0,k} \cdot n_{1,k}}}$$

O coeficiente correlação Phi varia entre -1 e 1. Quanto mais próximo estiver destes extremos, maior será a associação entre as variáveis. O sinal negativo da correlação significa que as variáveis variam em sentido contrário, isto é, que as categorias mais elevadas de uma variável estão associadas a categorias mais baixas da outra variável. Pelo contrário, o sinal positivo indica que as variáveis variam no mesmo sentido. Um valor próximo de zero é indicativo de não existência de relação entre as variáveis. Existem várias formulações para este coeficiente, sendo comumente aceita uma formulação onde o valor de Phi varia entre 0 e 1. Neste trabalho optou-se por se seguir a formulação apresentada, pois é coincidente com os *softwares* utilizados.

### **Coefficiente de correlação de contingência**

O coeficiente de contingência (C) permite correlacionar duas variáveis nominais, uma com  $n_{linhas}$  categorias e outra com  $n_{colunas}$  categorias, presentes em tabelas de contingência, não exigindo qualquer tipo de pressuposto relativamente à distribuição dos dados.

Este coeficiente pode ser calculado do seguinte modo (Conover, 1999):

$$C = \sqrt{\frac{\chi^2}{n + \chi^2}},$$

onde  $n$  representa o número de elementos na amostra e  $\chi^2$  a estatística de qui-quadrado:

$$\chi^2 = \sum_{i=1}^{n_{linhas}} \sum_{j=1}^{n_{colunas}} \frac{\left(n_{ij} - \frac{n_{i \cdot} \cdot n_{\cdot j}}{n}\right)^2}{\frac{n_{i \cdot} \cdot n_{\cdot j}}{n}}.$$

De notar que este coeficiente varia entre 0 e 1, nunca alcançando o valor máximo, mesmo que as variáveis estejam perfeitamente correlacionadas, uma vez que:

$$0 \leq C \leq \sqrt{\frac{(q-1)}{q}},$$

onde  $q = \min(n_{linhas}, n_{colunas})$ .

O valor nulo é indicativo de não existência de relação entre as variáveis.

### Coeficiente de correlação de Spearman

O coeficiente  $\rho$  de Spearman mede a intensidade da relação entre duas variáveis medidas numa escala pelo menos ordinal, usando para tal a ordem das observações. Deste modo, este coeficiente não é sensível a assimetrias na distribuição, nem à presença de outliers, não exigindo portanto que os dados sigam uma distribuição normal, no caso em que as variáveis em estudo são quantitativas. Este método revela-se eficaz em casos em que exista uma relação crescente ou decrescente entre as variáveis.

Podemos calcular o coeficiente de Spearman da seguinte forma (Maroco,2007):

$$\rho = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n^3 - n},$$

onde  $n$  é o número de pares  $(x_i, y_i)$  e  $d_i$  representa as diferença existentes entre as ordens dos valores correspondentes de  $(x_i, y_i)$ .

O coeficiente  $\rho$  de Spearman varia entre -1 e 1. Quanto mais próximo estiver destes extremos, maior será a associação entre as variáveis. O sinal negativo da correlação significa que as variáveis variam em sentido contrário, isto é, que as categorias mais elevadas de uma variável estão associadas a categorias mais baixas da outra variável. O sinal positivo indica que as variáveis variam no mesmo sentido. Um valor próximo de zero é indicativo de não existência de relação entre as variáveis.

## 2.2) Análise de Correspondências Múltiplas

A Análise de Correspondências (AC) é um método de análise multivariada com o objetivo de estudar a dependência entre os indivíduos e as categorias das variáveis, representados através de uma tabela de contingência. Este método baseia-se na decomposição do qui-quadrado de contingência, sendo o estudo da dependência realizado com base em representações gráficas (em que dois indivíduos ou duas categorias se assemelham tanto mais quanto mais próximos estiverem um do outro) e por parâmetros numéricos que permitem fazer a sua interpretação.

Ao longo deste trabalho iremos abordar, em particular, o método da Análise de Correspondências Múltiplas (ACM). A ACM representa um método de análise estatística de dados multivariados, que permite realizar comparações entre várias variáveis através da visualização e redução de dados contidos em tabelas



de grandes dimensões. Utilizando um método fatorial, hierarquiza os dados por ordem decrescente do grau de explicação do fenómeno em estudo e produz variáveis compósitas que resumem as relações existentes entre os atributos observados (Pestana e Gageiro, 2003).

A Análise de Correspondências Múltiplas é uma generalização da AC para o caso de tabelas com dimensão igual ou superior a três. Esta técnica é especialmente adaptada ao tratamento de um conjunto de respostas de um inquérito por questionário (no qual as perguntas representam as variáveis).

Poderemos resumir sumariamente as etapas da ACM em 3 grandes grupos: a quantificação das categorias, a seleção das dimensões e a representação gráfica das medidas de discriminação e respetiva interpretação.

### **Quantificação das categorias**

Numa primeira fase, e partir da matriz de *input* (matriz multidimensional na qual se dispõem  $n$  indivíduos linhas da matriz, caracterizados por  $m$  variáveis qualitativas – colunas), é criada uma matriz de presenças-ausências, matriz essa que será efetivamente utilizada pelo algoritmo da ACM (Carvalho, 2008).

Considerando a matriz de *input* sabemos que as categorias representam os objetos a ela associados e, por sua vez, os objetos são caracterizados por partilharem algumas categorias. Esta reciprocidade intrínseca à matriz é preservada no procedimento de transformação das categorias e dos objetos com o uso do princípio das médias recíprocas.

Para cada categoria de uma variável regista-se a presença de diferentes objetos. As categorias das variáveis efetuam uma partição entre os objetos definindo subgrupos que se caracterizam por partilharem certas categorias das diferentes variáveis em análise. A quantificação de cada categoria é representativa dos objetos que nela se inserem. Algebricamente a quantificação ( $Y$ ) das  $p$  categorias das  $m$  variáveis é dada por (Carvalho, 2008).

$$Y = D^{-1}.G'X \quad (1)$$

onde:

$Y$ : matriz de quantificações das categorias;

$D$ : matriz de frequências das  $p$  categorias (número total de categorias em análise);

$G$ : matriz de presenças-ausências de cada objeto nas categorias;

$X$ : matriz das coordenadas dos  $n$  objetos.

A expressão (1) evidencia que a quantificação das categorias ( $Y$ ) é igual à média das coordenadas dos

objetos que nelas se inserem ( $G'X$ ), ponderada pela frequência de ocorrência das categorias, a qual é retirada da matriz .

Tal como definida  $X$ , as coordenadas de cada objeto é proporcional à média das quantificações das categorias ( $Y$ ) associadas a cada objeto, isto é:

$$X \cong G.Y/m.$$

De modo a tornar esta proporcionalidade numa igualdade, bastará incluir um factor de proporcionalidade, cujo valor inverso será denominado por inércia ( $\lambda$ ). Assim, temos

$$X = G.Y/m \frac{1}{\lambda}.$$

### **Seleção das dimensões**

Com o objectivo de reduzir a dimensionalidade dos dados, será interessante poder visualizar os perfis, ainda que de forma aproximada, num espaço de poucas dimensões. Esta é a essência da Análise de Correspondências: a identificação de subespaços de poucas dimensões (preferencialmente com três ou menos) que contenham os perfis. Assim, é possível projetar no subespaço mencionado os perfis e observar as posições das suas projeções como uma aproximação às suas verdadeiras posições no espaço original de maior dimensionalidade. Na maioria dos casos necessitamos de pelo menos um plano para aproximar ou ajustar o espaço multidimensional da nuvem de perfis.

Neste sentido a ACM é classificada como um método de redução de dados, pois submete os dados a transformações matemáticas permitindo, assim, rever num espaço de menores dimensões a estrutura multifacetada e relacional do espaço de partida.

Assim, o primeiro passo consiste em determinar o número de dimensões a reter para a construção gráfica (e a partir das quais se vai basear a análise). A importância de cada uma das dimensões para explicar a variância dos dados de *input*, pode ser analisada através dos valores próprios e da inércia.

Os valores próprios quantificam a variância explicada por dimensão e é comum usar-se um critério de escolha das dimensões cujo valor próprio é superior a 1. A inércia varia entre 0 e 1 e quanto mais perto do limite superior mais variância é explicada por dimensão. Assim, a importância das dimensões está hierarquizada e vai diminuindo. É habitual privilegiar um número pequeno de dimensões, nas quais se registam os valores de inércia mais elevados.

É importante notar que é comum, no contexto da ACM, que os valores da inércia serem baixos, mesmo para as primeiras dimensões. Benzécri (1982) afirma que apesar dos valores da inércia (e por consequência a variância explicada) serem potencialmente baixos, não se deve cessar a interpretação da análise, sendo que a pouca representatividade numérica não deverá significar falta de qualidade na análise. Caso os perfis individuais não se afastem muito do perfil médio, os valores da inércia serão naturalmente baixos, porém não perdem qualquer tipo de relevância na interpretação por esse motivo.

O modelo de ACM admite no máximo  $c - v$  dimensões, onde  $c$  é o número de categorias presentes no estudo e  $v$  é o número de variáveis sem valores omissos. Caso todas as variáveis possuam valores omissos, o número máximo de dimensões é dado por  $c - 1$ . É com este valor máximo de dimensões que deverá ser iniciado o estudo, passando posteriormente para a seleção das dimensões relevantes para o modelo.

Uma vez determinado o número de dimensões a reter é comum fazer-se a denominação das dimensões encontradas com base nos valores relevantes para a dimensão das mesmas resultantes da matriz de discriminação.

## **Representação gráfica**

A representação gráfica das medidas de discriminação possibilita a observação da disposição das múltiplas variáveis nos planos definidos e a avaliação da relevância de cada uma delas nas dimensões. Quanto mais afastadas estiverem as variáveis da origem do gráfico e mais próximas de uma única dimensão, maior é o indício da presença de dimensões que envolvem traços de caracterização distintos. Por outro lado, variáveis próximas da origem correspondem a variáveis que não são diferenciadoras para as duas dimensões consideradas nesse plano.

A representação gráfica dos resultados da ACM permite visualizar no plano a nuvem de pontos das categorias das variáveis, assim como a nuvem de pontos dos objetos caracterizados pelas variáveis em causa. A nuvem de pontos dos objetos é formada a partir do conjunto dos perfis de cada linha. Da mesma forma, a nuvem de pontos das categorias é formada a partir do conjunto de perfis de cada coluna. A partir da leitura do plano é possível a identificação de associações entre categorias (proximidade das projeções).

## **Articulação da Análise de Correspondência com a Análise de Clusters**

A ACM permite, dispondo os indivíduos ou as categorias em planos, visualizar as relações que existem entre as múltiplas variáveis em estudo. Quando existe proximidade entre categorias de diferentes variáveis é induzida a presença de indivíduos que partilham tendencialmente as mesmas características, reflectindo a presença de grupos com perfis diferentes. Aplicando a Análise de Clusters e conhecendo-se a configuração topologia do espaço em análise é possível definir os grupos e saber o peso de cada um deles na amostra.

Para além de todas as combinações gráficas dimensionais possíveis, a ACM permite ainda a disposição das coordenadas das variáveis de cada dimensão, sendo que os valores das coordenadas são utilizados como variáveis de *input* (quantitativas) na Análise de Clusters.

Neste sentido, Carvalho (2008) refere “para efetuar a articulação ACM/Análise de Clusters usam-se como variáveis de *input* as coordenadas dos objetos (indivíduos) nas dimensões que sustentam o plano que, por sua vez, configura os grupos que se pretendem vir a definir” (...) “pode assim ver-se na articulação destes métodos de Análise de Dados uma estratégia para, a partir da configuração topológica, se passar à definição de tipologias”.

Assim, o interesse da combinação dos dois métodos justifica-se uma vez que pode resultar num aperfeiçoamento dos resultados da análise. Embora cada técnica possua particularidades e objetivos específicos de pesquisa, a associação de técnicas diferentes e complementares permite o refinamento das soluções e conclusões encontradas.

## 2.3) Análise Discriminante

A Análise Discriminante (AD) é um método estatístico multivariado cuja utilização visa, genericamente, identificar as características que distinguem os membros de um grupo dos de outro, de modo que, conhecidas as características de um novo indivíduo, se possa prever a que grupo pertence. (Maroco, 2007)

Esta técnica foi originalmente desenvolvida na Botânica tendo por objectivo fazer a distinção de grupos de plantas com base no tamanho e no tipo de folhas o que tornaria possível, posteriormente classificar as novas espécies encontradas. Em 1936, Fisher foi o responsável pelo desenvolvimento da análise para dois grupos.

Por via a escolher quais as variáveis que distinguem os grupos em análise, são criadas funções discriminantes, provenientes de combinações lineares das variáveis iniciais, que maximizam as diferenças entre as médias dos grupos e minimizam as classificações incorrectas dos casos nos grupos.

Um dos métodos utilizados para descobrir as variáveis que discriminam os grupos é a estatística  $\Lambda$  de Wilks, que se trata de uma medida inversa do grau de diferenciação entre os grupos: quanto menor o seu valor, maior o grau de diferenciação entre estes. Esta estatística é dada por (Maroco,2007):

$$\Lambda = \frac{|W|}{|T|} = \frac{|W|}{|B + W|}$$

onde  $|W|$  representa a medida da variabilidade dentro dos grupos,  $|B|$  representa a medida da variabilidade entre os grupos e  $|T|$  a medida da variabilidade total. Quanto maior for a semelhança entre os dois determinantes, menores serão as diferenças entre os grupos e mais próximo o Lambda de Wilks estará de 1.

A função discriminante linear ou de Fisher permite caracterizar cada um dos grupos através de uma combinação linear das variáveis iniciais, permitindo verificar quais as variáveis mais importantes para cada um dos grupos, e assim determinar a existência de diferenças ou não entre grupos. Deste modo é possível caracterizar cada um dos  $k$  grupos da seguinte forma:

$$GR_k = \lambda_{ik} var_i + \lambda_{jk} var_j + \dots + \lambda_{nk} var_n$$

## 2.4) Regressão Logística

A Regressão Logística (RL) é um modelo particular da classe dos Modelos Lineares Generalizados (GLM) nos quais a variável dependente é dicotómica e onde função de ligação é do tipo *logit*. À semelhança de outros GLM, o objetivo da regressão logística é explorar a relação entre uma ou mais variáveis explicativas (ou independentes) e uma variável resposta (ou dependente).

Assim, a formulação do modelo de regressão logística é dada por (Homer & Lesmeshow, 2000):

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x + \dots + \beta_k x}}{1 + e^{\beta_0 + \beta_1 x + \dots + \beta_k x}},$$

onde os parâmetros  $\beta_0, \beta_1, \dots, \beta_k$  são estimados pelo método de máxima verosimilhança.

Utilizando as propriedades algébricas da função anterior e fazendo, a transformação logaritmica:

$$g(x) = \ln\left(\frac{\pi(x)}{1 - \pi(x)}\right),$$

temos a função logística linearizada (ou função *logit* de probabilidade)

$$g(x) = \beta_0 + \beta_1 x + \dots + \beta_k x,$$

onde  $g(x) \in \mathbb{R}$ .

Uma das primeiras etapas a realizar no uso regressão logística passa por determinar o modelo que melhor se ajusta aos dados em análise, com o intuito de se obter um modelo que permita descrever a relação entre a variável resultado e um conjunto de variáveis independentes (Braga, 1994). Neste sentido, há que determinar quais as variáveis que são relevantes para o modelo de regressão logística (Maroco, 2007:703-705).

Quando são cumpridos os pressupostos geralmente associados aos GLM, nomeadamente dada a estrutura aleatória dos dados, a selecção das variáveis significativas para o modelo é feita com base no teste de Wald ou o teste da razão de verosimilhança (Homer e Lesmeshow, 2000).

No caso da violação de pressupostos base dos pressupostos dos GLM (Maroco, 2007:685-703). é ainda mais importante a validação do modelo, utilizando como medida a qualidade do ajustamento e o poder discriminante do modelo. Assim, de modo a avaliar o poder discriminante de um modelo de regressão logística, utiliza-se a curva ROC, gráfico no qual, o eixo vertical apresenta a percentagem dos verdadeiro-positivo (a sensibilidade) e no eixo horizontal apresenta os verdadeiros-negativos (especificidade). A curva

ROC pode ser feita por meio de um gráfico, que permite estudar a os valores da sensibilidade e especificidade, para diferentes valores de corte. A sensibilidade é definida como a percentagem de casos que o modelo forneceu um resultado verdadeiro-positivo, enquanto a especificidade é definida como a percentagem de vezes que o modelo forneceu um verdadeiro-negativo. Deste modo e aplicando estas métricas aos dados iniciais, poderemos obter uma percepção da qualidade do ajustamento.

Outro indicador importante é dado pela curva ROC é a área abaixo da curva (AUC). Com base neste valor é possível determinar o quão discriminativo é o modelo encontrado. É possível interpretar os valores de AUC com base na Tabela 3 (Maroco,2007).

**Tabela 3:** Poder discriminativo do modelo de regressão logística com base na área abaixo da curva ROC

Área ROC (AUC)	Poder discriminativo
0,5	Sem poder discriminativo
]0,5; 0,7[	Discriminação fraca
[0,7; 0,8[	Discriminação aceitável
[0,8; 0,9[	Discriminação boa
≥0,9	Discriminação excepcional

Através dos valores dos coeficientes do modelo de regressão logística obtidos, é ainda possível retirar conclusões sobre as categorias da variável dependente. Esta interpretação é efectuada através do cálculo da razão das chances (*odds ratio*), isto é a razão entre a possibilidade de pertencer a uma categoria da variável dependente face a outra. Desta forma, pretendemos comparar se a possibilidade de um determinado evento é a mesma para dois grupos. Para um modelo de regressão logística, sem interações entre as variáveis independentes, para uma dada variável o *odds ratio* é dado por (Homer & Lesmeshow, 2000):

$$\widehat{OR} = e^{\hat{\beta}}.$$

### 3) Metodologia

Depois da revisão da literatura sobre o aparecimento e evolução do *e-commerce*, assim como dos principais modelos estatísticos que serão utilizados no decorrer deste trabalho, inicia-se este terceiro capítulo com a apresentação da metodologia do estudo: as hipóteses que conduziram a este trabalho, a caracterização do universo, a definição da amostra e os procedimentos para a construção do questionário e recolha dos dados.

#### 3.1) Hipóteses Teóricas em Estudo

A realização deste estudo visa compreender os hábitos de consumo *online*, sendo que para tal, foram formuladas as seguintes hipóteses que se pretendem ser estudadas:

**Hipótese 1)** Não existem diferenças entre géneros no que diz respeito a comprar (ou não) *online*;

**Hipótese 2)** Existem diferenças entre os produtos comprados por homens e por mulheres;

**Hipótese 3)** Existem diferenças entre os serviços que homens e mulheres aderem *online*;

**Hipótese 4)** As inquiridas do sexo feminino tem maior tendência a comprar por impulso relativamente aos homens;

**Hipótese 5)** Existem diferenças entre as vantagens da compra *online* identificadas por compradores e não compradores;

**Hipótese 6)** Existem diferenças entre as vantagens da compra *online* identificadas por homens e mulheres;

**Hipótese 7)** Existem diferenças entre as desvantagens da compra *online* identificadas por compradores e não compradores;

**Hipótese 8)** Existem diferenças entre as desvantagens da compra *online* identificadas por homens e mulheres;

**Hipótese 9)** Existem diferenças entre os meios de comunicação digital preferidos por homens e mulheres;

**Hipótese 10)** Existem diferenças entre os meios de comunicação digital preferidos por compradores e não compradores;

Para além das hipóteses formuladas serão abordadas outras questões, colocadas ao longo do inquérito, por via a perceber melhor os perfis de consumo da amostra em estudo.



### 3.2) Caracterização do Universo do estudo e definição da amostra

Segundo Hill e Hill (2002), ao conjunto total dos casos (conjunto de entidades que fornecem dados na forma de observações ou medidas dos valores de uma ou várias variáveis) sobre os quais se pretendem retirar conclusões dá-se o nome de População ou Universo. Uma amostra é um subconjunto de indivíduos da população, que se analisa com o objetivo de retirar conclusões sobre essa população de onde foi recolhida (Hill e Hill, 2002). Como veremos, ao longo deste estudo, nem sempre é possível inferir resultados para a população, devido à não representatividade da amostra em causa, o que pode enviesar as conclusões. Nesses casos a análise deve-se limitar-se à interpretação dos resultados do grupo concreto dos indivíduos inquiridos.

Os processos a adotar na recolha da amostra agrupam-se essencialmente em duas categorias (Vicente *et al.* 2001): aleatórios ou probabilísticos – quando cada indivíduo da população tem uma probabilidade conhecida de fazer parte da amostra; não aleatórios ou dirigidos – quando há indivíduos da população que não têm possibilidades de ser escolhidos. A opção por um processo de amostragem probabilístico carece da necessidade existência de uma base de amostragem, ou seja, uma listagem da totalidade dos indivíduos ou unidades amostrais da população.

Neste trabalho o universo em estudo é constituído pelos utilizadores da *internet*, com pelo menos 18 anos. A dimensão desta população é desconhecida e, conseqüentemente, não existe uma base de amostragem da mesma. Assim sendo, para a recolha da amostra optou-se por um processo de amostragem não probabilístico. Tendo em conta o objetivo do estudo e os meios disponíveis, optou-se pelo processo de amostragem bola de neve. Esta técnica é, tipicamente, utilizada em populações raras, de dimensão desconhecida ou cujos membros não são possíveis de identificar (Coleman, 1958; Goodman, 1961; Spreen, 1992). Baseia-se na escolha, por conveniência ou não, de um conjunto inicial de indivíduos da população aos quais é pedida a participação no estudo e ainda que estes estimulem a participação no estudo de indivíduos do universo por si escolhidos. A analogia com a bola de neve é facilmente compreendida uma vez que o expectável é que desde o momento em que o estudo é lançado pelo utilizador, o número de participantes cresça a cada iteração. Apesar da facilidade com que permite a divulgação do questionário, a amostragem em bola de neve pode apresentar constrangimentos ao nível da representatividade e enviesamento da amostra: os inquiridos têm tendência a partilhar o estudo com amigos ou pessoas mais próximas, o que poderá conduzir a uma amostra cujos elementos possuem uma certa homogeneidade entre si, compartilhando traços e características, podendo até levar a amostra a ser apenas um subgrupo específico do universo de estudo (Faugier e Sargeant, 1997).

### 3.3) Construção do Questionário e Recolha da Informação

Um questionário é definido por Hill e Hill (2002) como uma técnica de investigação composta por um número grande ou pequeno de questões apresentadas por escrito que tem por objetivo propiciar determinado conhecimento ao pesquisador; é um processo rigoroso e dele dependem a medição das variáveis da investigação

Segundo Hoz (1985) é um instrumento para recolha de dados constituído por um conjunto mais ou menos amplo de perguntas e questões que se consideram relevantes de acordo com as características e a dimensão do que se deseja observar.

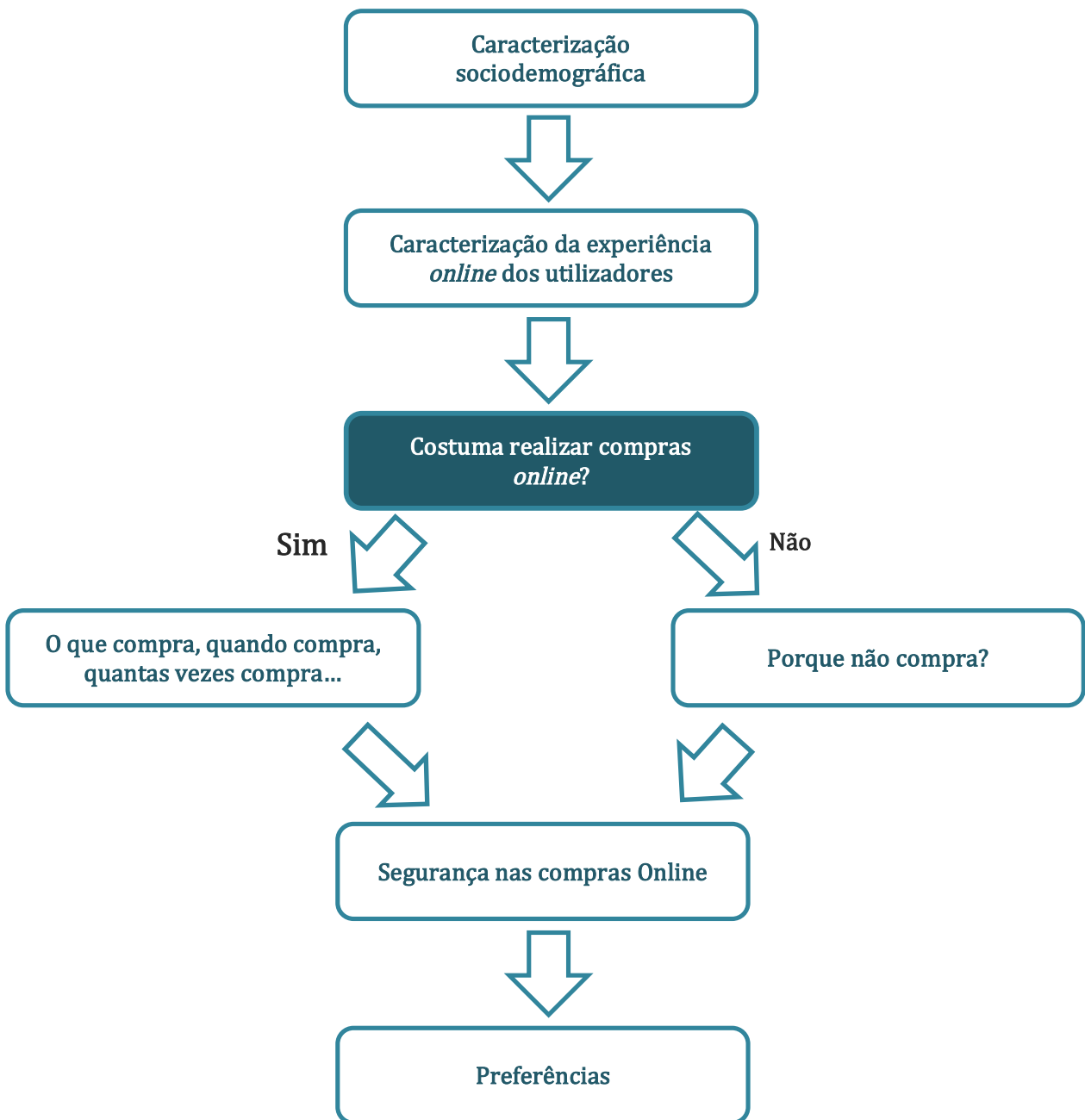
O encadeamento metodológico sobre o qual foi construído o questionário (representado na Figura 2, que tem como principal “pergunta-chave”: “**Costuma realizar compras *Online*?**” permitiu definir um caminho distinto conforme a resposta dos utilizadores, tentando explorar as razões que os levam a comprar ou não *online*, estudando para ambos os grupos quais as suas motivações, as vantagens e receios ao comprar *online* assim como outros aspetos relacionados com os seus hábitos de consumo.

No período de conceção do questionário foram aplicadas duas versões preliminares do mesmo, a uma amostra de indivíduos, com o objetivo de identificar possíveis as questões que careciam de reformulação. Com base nos comentários obtidos, foi elaborada a versão final que se apresenta no Anexo 1.

Recorreu-se à ferramenta de *Formulários* do *Google Drive* para construir o questionário em formato eletrónico de modo a que fosse, por um lado, facilmente partilhável e inteligível e, por outro lado, um método simples e fiável de integração dos dados, diminuindo assim os erros de compilação pelo utilizador. Esta ferramenta foi também bastante importante para a diminuição dos erros por parte dos utilizadores, impedindo, por exemplo que utilizadores que não costumam realizar compras *online* respondessem a questões destinadas apenas aos que têm esse hábito. Uma vez que os métodos escolhidos para partilha do questionário foram o *email* e redes sociais, todos os indivíduos que não utilizem este tipo de ferramentas foram automaticamente excluídos do estudo.

Após a recolha da amostra, a codificação dos dados e posterior tratamento estatístico foram realizados nos programas Excel, IBM SPSS Statistics, versão 22, e R Project.

Figura 2: Fluxo do questionário aplicado



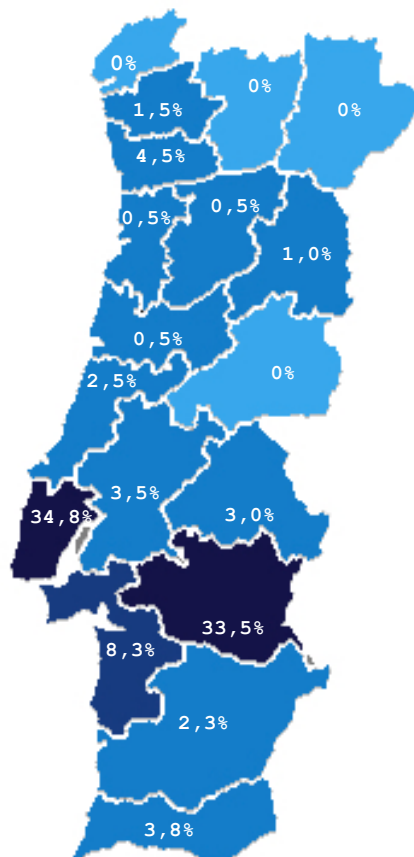
## 4) Apresentação e Discussão dos Resultados

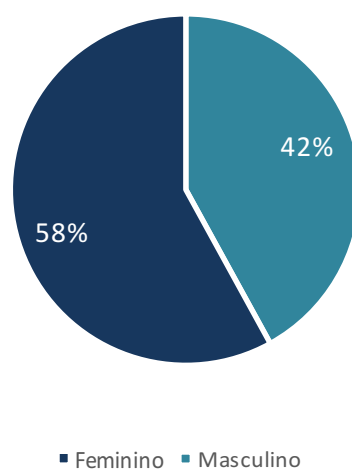
Neste capítulo pretende-se caracterizar os hábitos de *e-commerce* dos inquiridos, apresentando para tal os resultados e a análise dos resultados ao inquérito aplicado. Seguindo a ordem natural do inquérito, é efetuada uma análise descritiva dos dados, sendo estudadas cada uma das hipóteses colocadas de acordo com os métodos de se consideraram mais relevantes. Para os temas relacionados com as vantagens e desvantagens da compra *online* assim como dos meios de publicidade mais eficazes, serão exploradas três abordagens distintas (Análise de Correspondências Múltiplas, Análise Discriminante com particular incidência na função de Fisher e Regressão Logística) com o objetivo de comparar os resultados e as variáveis escolhidas entre os vários métodos.

### 4.1) Análise sociodemográfica da amostra

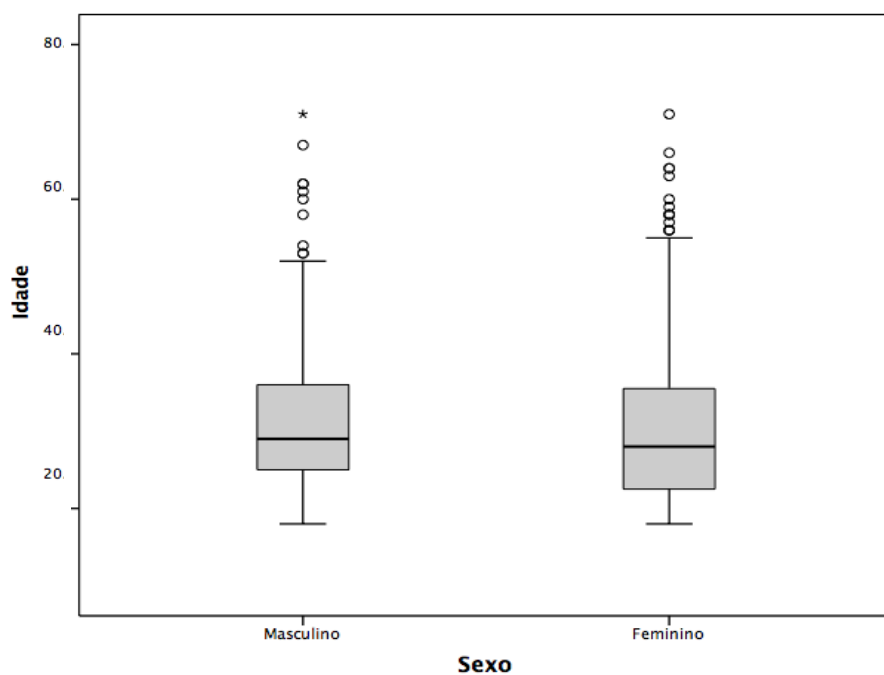
Participaram no estudo 400 indivíduos, residentes em 14 dos 18 distritos de Portugal, sendo o distrito mais representado Lisboa (34,8%), seguido de Évora (33,5%) e Setúbal (8,3%) - Figura 3.

**Figura 3:** Percentagem de indivíduos da amostra divididos pelos distritos de Portugal



**Gráfico 3:** Sexo dos inquiridos

A amostra é constituída por indivíduos entre os 18 e os 71 anos, sendo a maioria do sexo feminino (Gráfico 3). Os homens do estudo apresentam uma idade média ( $\bar{x} = 31,6$  anos;  $s = 9,8$ ) superior à das mulheres ( $\bar{x} = 30,8$  anos;  $s = 11,4$ ). Para ambos os sexos observa-se uma assimetria positiva na distribuição das idades, indicando assim uma maior frequência de inquiridos com idades mais jovens (Gráfico 4). Observa-se ainda a existência de *outliers* moderados superiores quer na distribuição das idades dos homens quer na das mulheres e ainda um *outlier* severo nos homens.

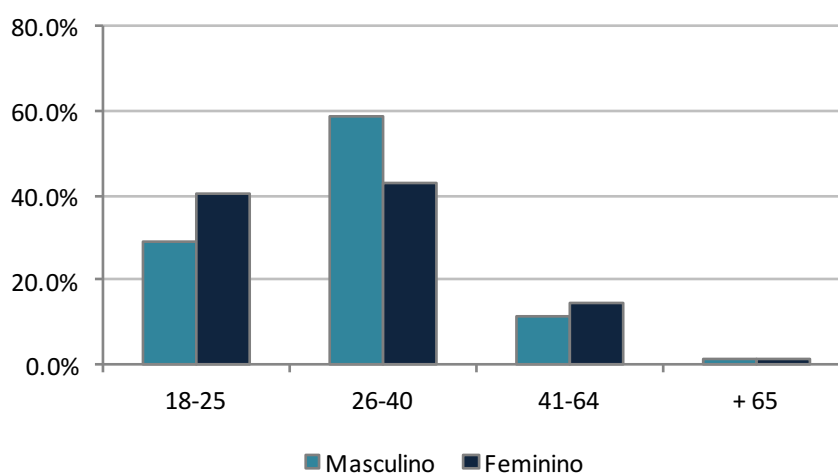
**Gráfico 4:** Caixa de bigodes idade por sexo

Por conveniência de tratamento dos dados foi feita uma categorização da variável idade em 4 classes, em função da sua familiaridade com a *internet*:

- **Jovens-adultos** [18; 26[: faixa etária que se caracteriza pela entrada na vida adulta e a que geralmente corresponde um menor poder financeiro, mas a que está também associado um maior conhecimento informático, uma vez que esta é uma geração nativa em questões relacionadas com a Internet;
- **Adultos** [26; 41[: faixa etária que corresponde a uma geração que cresceu em paralelo com o mundo informático e que possui maior poder de compra face aos jovens-adultos;
- **Adultos Maduros** [41; 65[: representantes de uma geração que apenas contactou com a Internet na vida adulta;
- **Séniore**s [65; + ∞[: faixa etária que representa os indivíduos que durante grande parte da sua vida adulta não tiveram acesso à Internet.

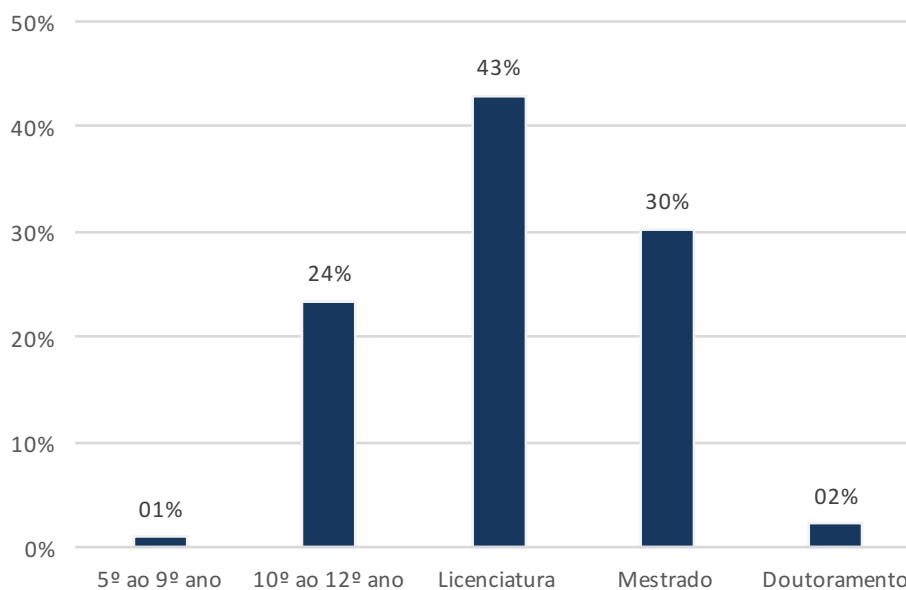
Na amostra em estudo, predominam os adultos e somente 1% tem 65 ou mais anos, ou seja, são séniore (Gráfico 5). É ainda possível verificar que entre as mulheres a percentagem de jovem-adultas é semelhante à das adultas, enquanto entre os homens predominam os adultos.

**Gráfico 5:** Distribuição dos indivíduos nas classes de idade, por sexo



Mais de 75% dos inquiridos possui grau de instrução universitária (Licenciatura, Mestrado ou Doutoramento), o que denota já nesta fase alguns problemas associados à representatividade dos dados por via da opção tomada face ao método de amostragem utilizado: de acordo com os CENSOS 2011, apenas 11,9% dos Portugueses tem formação superior.

**Gráfico 6:** Distribuição dos indivíduos por grau de instrução



Cerca de 66,8% dos inquiridos (267 indivíduos) são solteiros, 30,0% são casados ou vivem em união de facto, 2,5% são divorciados e os restantes viúvos. A maioria (75,0%) dos inquiridos não tem filhos, sendo que 14,8% tem um filho, 8,0% tem 2 filhos e apenas 2,3% tem 3 filhos ou mais.

## 4.2) Experiência *Online* dos Inquiridos

Uma das questões colocadas aos inquiridos foi acerca da sua auto-perceção do quão à vontade estavam com o uso da internet, sendo que 89% deram a si próprios uma avaliação positiva (nota 6 ou superior, numa escala onde 1 representava utilizadores “nada à vontade” e 10 utilizadores “totalmente à vontade”). Este valor tão expressivo pode ser explicável pela frequência de utilização da internet por parte dos inquiridos: 97,5% afirmou utilizar a internet **todos os dias** ou mesmo **constantemente**.

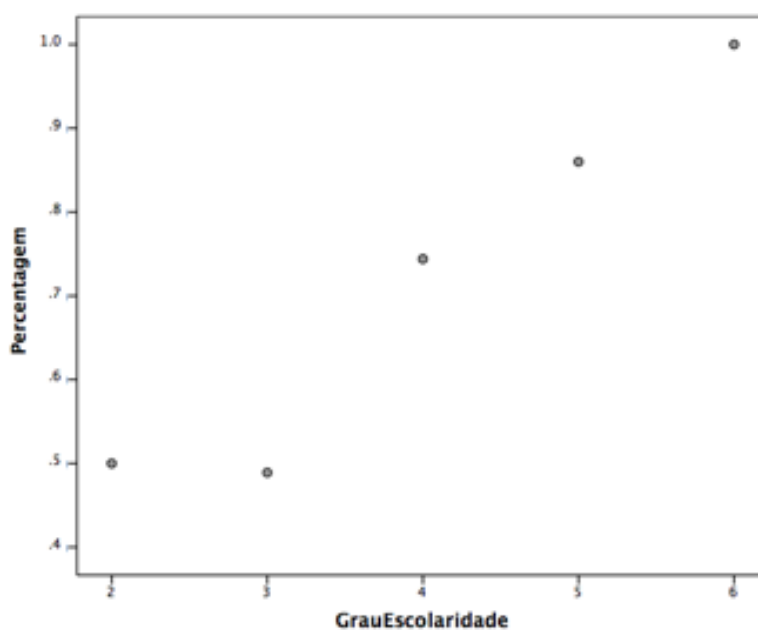
Da análise da Tabela 4 verifica-se que cerca de 3 em cada 4 inquiridos (72%) respondeu que **costuma realizar compras online**, sendo que entre os homens a proporção dos que compram online é superior à observada entre as mulheres (85,1% contra 62,9%). Esta associação entre o comprar ou não *online* e o sexo é fraca ( $\Phi = 0,245$ ;  $C = 0,238$ ).

**Tabela 4:** Hábito de comprar *online* por sexo

		Sexo		Total
		Feminino	Masculino	
Costuma realizar	Não	86	25	111
compras Online?	Sim	146	143	289
Total		232	168	400

Verifica-se que existe também uma associação fraca entre o comprar ou não *online* e a região de residência dos inquiridos ( $\Phi = 0,277$ ), sendo na região de Lisboa aquela onde se regista uma maior diferença entre a proporção de compradores e a de não compradores, comparativamente às restantes regiões.

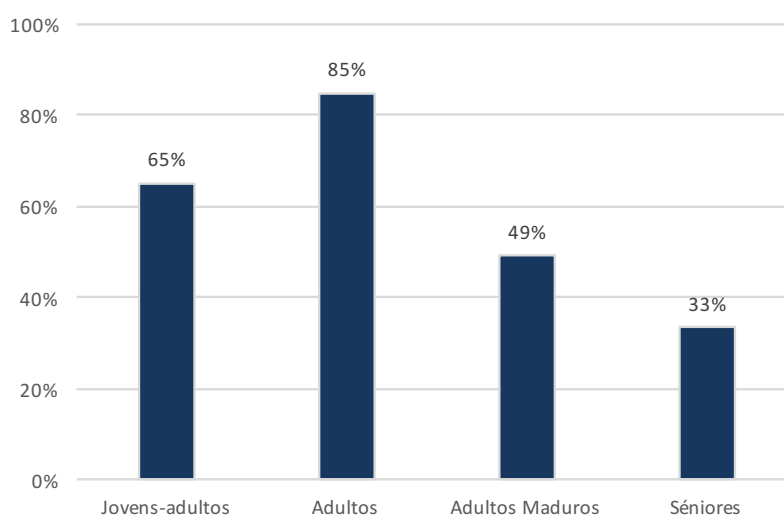
Considerando a percentagem de inquiridos que costuma realizar compras *online* por grau de habilitações académicas, observa-se que existe uma associação positiva forte entre estas duas variáveis ( $\rho = 0,90$ ; Gráfico 7), o que indica que quanto maior o grau de habilitações académicas dos indivíduos maior é a proporção de inquiridos que costuma efetuar compras *online*.

**Gráfico 7:** Gráfico de dispersão da percentagem de indivíduos que compra *online* em função do grau de instrução



Tendo em conta a variável idade categorizada, da forma já referida anteriormente, foram os jovens-adultos inquiridos os que mais compraram *online*, tendo 84,8% deles referido ter comprado *online* (Gráfico 8), seguindo-se os jovens tendo 65,0% deles já feito transações *online*. Estas percentagens são explicadas por se tratarem de faixas jovens da população, mais habituadas a um mundo digital. Também é natural pensar-se que por via de um maior poder de compra e maior independência os inquiridos com idade entre 26 e 39 anos compreem mais que os inquiridos com idade igual a inferior a 25 anos.

**Gráfico 8:** Percentagem de compradores *online* em cada classe etária



### 4.3) Compradores *Online*

Nesta secção serão apenas tidos em consideração os indivíduos que responderam “Sim” à questão “**Costuma realizar compras *online*?**”, ou seja, 72% dos inquiridos

Tal como dita o senso comum, o estudo revelou diferenças claras entre os hábitos de compra *online* entre os inquiridos por sexo. À primeira vista, é fácil verificar que os homens têm uma apetência muito mais forte por comprar produtos relacionados com tecnologia, enquanto as mulheres se diferenciam por comprar mais roupa e acessórios pessoais (Gráfico 9).

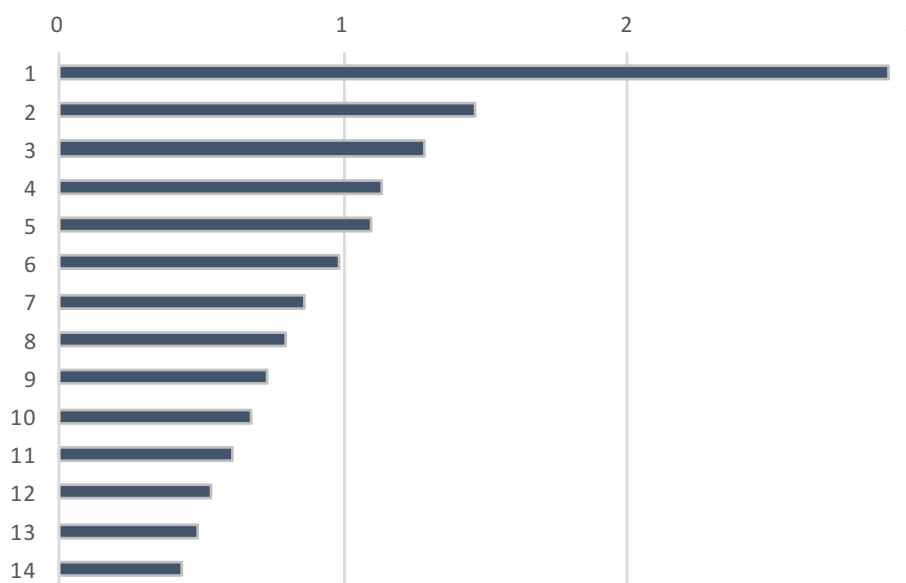
**Gráfico 9: Produtos comprados por mulheres e homens**

Por forma a completar a análise da hipótese 2 (“Existem diferenças entre os produtos comprados por homens e por mulheres”), iremos realizar uma análise multivariada dos produtos adquiridos *online*. Serão usadas três técnicas diferentes (a Análise de Correspondências Múltipla, a Análise Discriminante e a Regressão Logística) com o objectivo de ilustrar as diferenças/semelhanças entre estas técnicas.

### **Análise de correspondências múltiplas:**

Recorremos à ACM para realizar uma análise conjunta dos produtos comprados *online*, na averiguação de diferenças entre géneros. Começou-se por se definir o número máximo de dimensões que permitem representar este conjunto de dados, i.e., 14 (Anexo 2). Em função dos resultados se procurou conjugar o decaimento dos valores próprios (

Gráfico 10) e da análise das medidas de discriminação obtidas (Tabela 5) para fazer a seleção do número de dimensões para o modelo. Através destes métodos foram seleccionadas 4 dimensões, uma vez que a partir da 5ª dimensão, as variáveis relevantes para cada dimensão não adicionam nada de novo face às anteriores. Assim, o modelo escolhido representa 41,0% da variância total das variáveis.

**Gráfico 10:** Representação da variância das dimensões associadas aos produtos comprados *online***Tabela 5:** Medidas de discriminação associadas aos produtos comprados *online*

	Dimension				Mean
	1	2	3	4	
Bilhetes para Concertos/Espétaculos	<b>0,340</b>	0,147	0,001	0,012	0,125
Acessórios Pessoais (Relógios, Malas, Perfumes...)	0,131	<b>0,211</b>	<b>0,278</b>	0,000	0,155
Automóvel ou outro veículo motorizado	0,015	0,008	0,001	0,045	0,017
Brinquedos	<b>0,286</b>	0,048	0,036	0,116	0,121
Livros	0,066	0,137	<b>0,272</b>	0,021	0,124
Experiências / Vouchers	<b>0,292</b>	<b>0,244</b>	0,093	0,008	0,159
Filmes/Séries	0,150	0,032	<b>0,293</b>	0,001	0,119
Roupa	0,101	<b>0,244</b>	<b>0,231</b>	0,021	0,149
Tecnologia	<b>0,202</b>	0,020	0,033	<b>0,293</b>	0,137
Produtos para o Lar e Decoração	<b>0,376</b>	0,075	0,023	0,038	0,128
Refeições	<b>0,339</b>	0,009	0,001	0,182	0,133
Viagens	<b>0,323</b>	<b>0,271</b>	0,000	0,003	0,149
Supermercado	<b>0,275</b>	0,003	0,018	<b>0,308</b>	0,151
Outros	0,016	0,014	0,000	0,088	0,030
Sexo <sup>a</sup>	0,028	0,009	<b>0,109</b>	0,017	0,041
Active Total	2,911	1,464	1,280	1,135	1,697
% of Variance	20,793	10,454	9,142	8,106	12,124

Com base nos valores das medidas de discriminação e considerando as variáveis que mais contribuem para a construção de cada uma das dimensões, passaremos a denominar cada uma das destas do seguinte modo:

- Dimensão 1: “**Generalistas**” (produtos: “Bilhetes para Concertos/Espétaculos”, “Brinquedos”, “Experiências / Vouchers”, “Tecnologia”, “Produtos para o Lar e Decoração”, “Refeições”, “Viagens” e “Supermercado”)
- Dimensão 2: “**Actividades e moda**” (produtos: “Acessórios Pessoais (Relógios, Malas, Perfumes...)”, “Experiências / Vouchers”, “Roupa” e “Viagens”)
- Dimensão 3: “**Lazer e moda**” (produtos: “Acessórios Pessoais (Relógios, Malas, Perfumes...)”, “Livros”, “Filmes / Séries” e “Roupa”)
- Dimensão 4: “**Tecnologia e supermercado**” (produtos: “Tecnologia” e “Supermercado”)

Através Tabela 5, podemos ainda verificar que o valores das medidas de discriminação da variável sexo, da dimensão 3 não é tão próximo de zero, o que indicia que podem existir de diferenças entre os produtos comprados *online* entre géneros.

Na Tabela 6 apresenta-se a forma como as categorias das várias variáveis mais relevantes para cada uma das dimensões se posicionam nos dois eixos das mesmas.

**Tabela 6:** Posicionamento das variáveis mais relevantes para a ACM dos produtos comprados *online*

	Quantificações negativas	Quantificações positivas
<u>Generalistas</u> (Dimensão 1)	Bilhetes para Concertos/Espétaculos_Não Brinquedos_Não Experiências / Vouchers_Não Tecnologia_Não Produtos para o Lar e Decoração_Não Refeições_Não Viagens_Não Supermercado_Não	Bilhetes para Concertos/Espétaculos_Sim Brinquedos_Sim Experiências / Vouchers_Sim Tecnologia_Sim Produtos para o Lar e Decoração_Sim Refeições_Sim Viagens_Sim Supermercado_Sim
<u>Actividades e moda</u> (Dimensão 2)	Acessórios Pessoais (Relógios, Malas, Perfumes...)_Sim Experiências / Vouchers_Não Roupa_Sim Viagens_Não	Acessórios Pessoais (Relógios, Malas, Perfumes...)_Não Experiências / Vouchers_Sim Roupa_Não Viagens_Sim
<u>Lazer e moda</u> (Dimensão 3)	Acessórios Pessoais (Relógios, Malas, Perfumes...)_Sim Livros_Não Filmes/Séries_Não Roupa_Sim	Acessórios Pessoais (Relógios, Malas, Perfumes...)_Não Livros_Sim Filmes/Séries_Sim Roupa_Não
<u>Tecnologia e supermercado</u> (Dimensão 4)	Tecnologia_Não Supermercado_Não	Tecnologia_Sim Supermercado_Sim

Assim, verifica-se que a dimensão **Generalista** associa os indivíduos que procedem de igual modo à compra *online* de Bilhetes para Concertos/Espétaculos, Brinquedos, Acessórios Pessoais, Experiências / Vouchers, Tecnologia, Produtos para o Lar e Decoração, Refeições, Viagens e Supermercado.

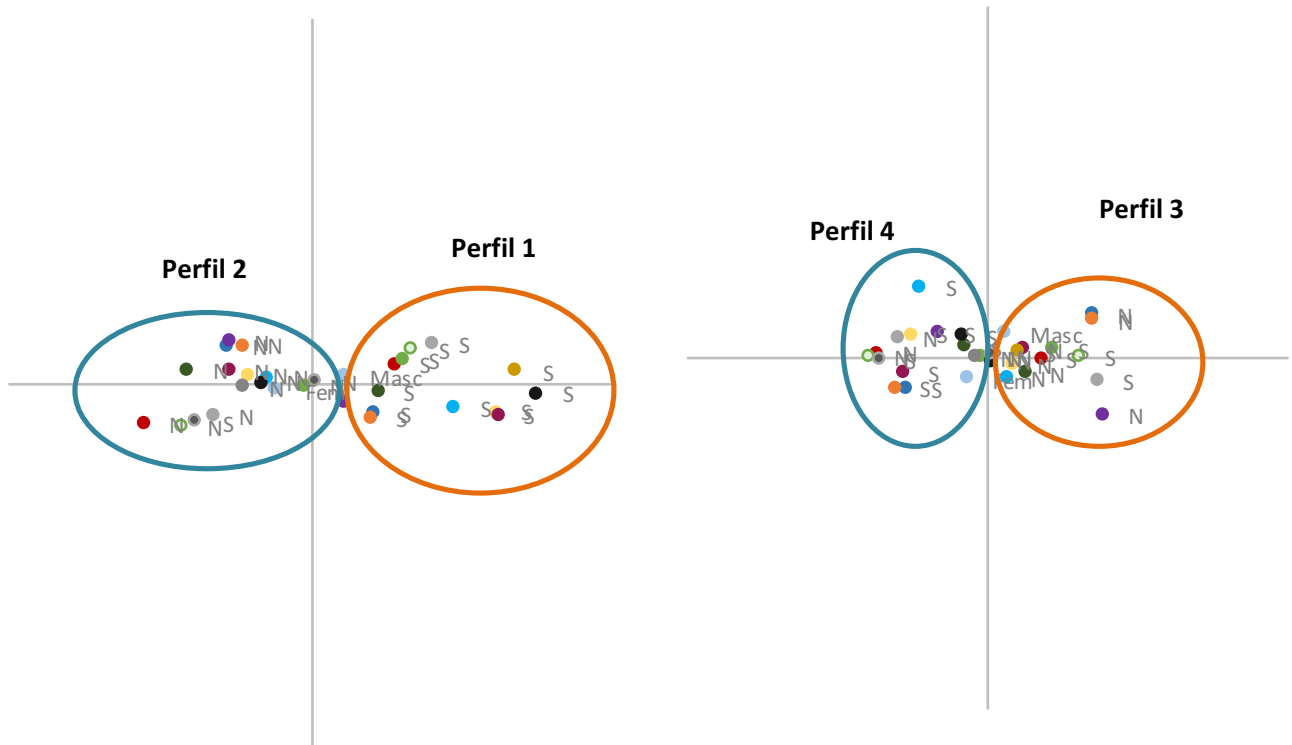
A dimensão de **Actividades e moda** caracteriza-se por associar os indivíduos que são opostos relativamente a comprar *online* produtos mais relacionados com actividades (Experiências / Vouchers e Viagens) aos indivíduos que compram produtos relacionados moda pessoal (Acessórios Pessoais (Relógios, Malas, Perfumes...) e Roupa).

A dimensão de **Lazer e moda** caracteriza-se por associar os indivíduos que são opostos relativamente a comprar *online* produtos mais relacionados com moda (Acessórios Pessoais (Relógios, Malas, Perfumes...) e Roupas) aos indivíduos que compram produtos relacionados com lazer (Livros e Filmes/Séries).

A dimensão de **Tecnologia e supermercado** caracteriza-se por associar os indivíduos que são concordam relativamente a comprar *online* tecnologia e produtos de supermercado.

**Gráfico 11:** Coordenadas dos centroides da ACM para os produtos comprados *Online*:  
**Dimensão 1 vs Dimensão 2**

**Gráfico 12:** Coordenadas dos centroides da ACM para os produtos comprados *Online*:  
**Dimensão 2 vs Dimensão 3**



- Bilhetes para Concertos/Espétaculos
- Acessórios Pessoais (Relógios)
- Automóvel ou outro veículo motorizado
- Brinquedos
- Livros
- Experiências / Vouchers
- Filmes/Séries
- Roupas
- Tecnologia
- Produtos para o Lar e Decoração
- Refeições
- Viagens
- Supermercado
- Outros
- Sexo

**Gráfico 13:** Coordenadas dos centroides da ACM para os produtos comprados *Online*:  
**Dimensão 3 vs Dimensão 4**



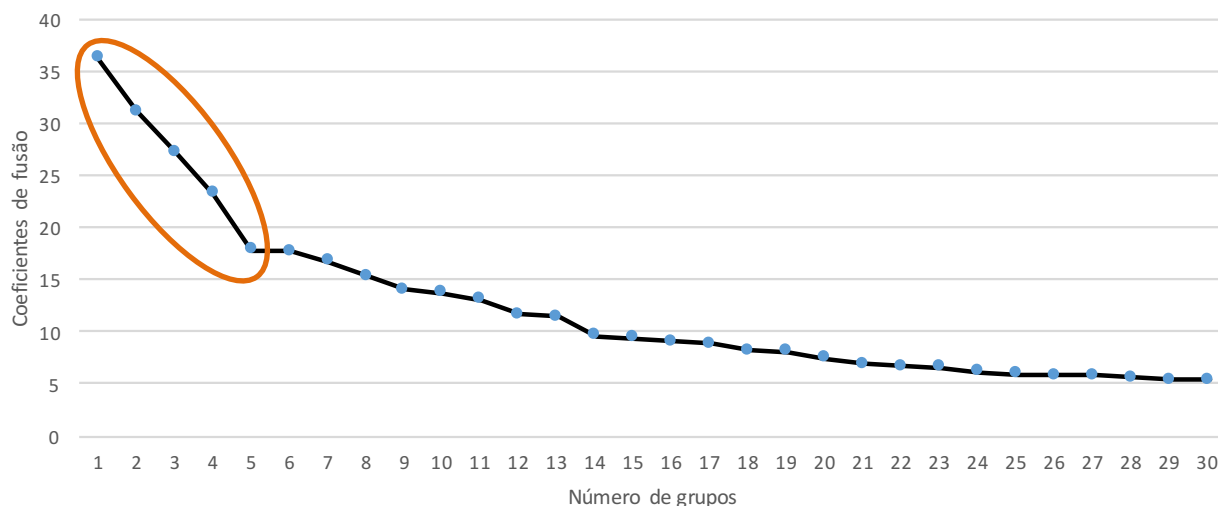
Da análise dos gráficos com a representação do espaço dos produtos (Gráfico 11-Gráfico 13), tendo em conta as variáveis que mais discriminam as respectivas dimensões, parecem existir 5 perfis de compradores:

- 1) Indivíduos que compram *online* quase todo o tipo de produtos;
- 2) Indivíduos que compram *online* produtos não diferenciados;
- 3) Indivíduos que compram essencialmente os produtos associados às actividades;
- 4) Indivíduos que compram produtos da área da moda;
- 5) Indivíduos que compram produtos ligados à área da tecnologia e audiovisual e não compram produtos associados à moda.

Destaca-se ainda que a coordenada associada ao sexo masculino está mais próxima dos perfis 2 e 5, o que indica que podem existir algumas diferenças entre os homens e as mulheres no que diz respeito aos produtos comprados *online*.

Posteriormente, realizou-se uma análise de *clusters* das coordenadas dos indivíduos nas dimensões retidas de modo a identificar grupos que correspondessem aos perfis encontrados. Numa fase inicial optou-se pela análise de *clusters* hierárquica para validar se o número de grupos correspondia ao número de perfis, e de seguida, realizou-se uma análise de *clusters* não hierárquica que otimiza a distância de cada indivíduo ao centroide.

**Gráfico 14:** Coeficientes de fusão da análise de *clusters* associada aos produtos comprados *online* – critério vizinho mais afastado



Da análise do gráfico dos coeficientes de fusão, e considerando que existem existe uma mudança de declive no 5º agrupamento, conclui-se que indicações dadas pelos com a representação do espaço dos produtos que também aqui há uma indicação para que existam 5 grupos de indivíduos, cujas principais características são (Tabela 7):

- O grupo 1 (15% dos compradores) é composto pelos indivíduos que compram quase todo o tipo de produtos e se dividem na compra dos brinquedos e dos filmes/séries. Este grupo é maioritariamente composto por homens;
- O grupo 2 (24% dos compradores) identifica o segmento dos compradores de produtos não diferenciados e que se dividem na compra de produtos de moda;
- O grupo 3 (22% dos compradores) refere um segmento mais consumidor de atividades;



- O grupo 4 (22% dos compradores) representa o segmento dos consumidores de produtos ligados à moda;
- O grupo 5 (17% dos compradores) representa um segmento de compradores de tecnologia, sendo na sua generalidade constituído por homens.

**Tabela 7:** Caracterização dos 2 grupos identificados, relativamente aos às vantagens da compra *online* identificadas pelos inquiridos.

	Categoria	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Total
Bilhetes para Concertos/Espétaculos	Não	4,5%	<b>73,5%</b>	22,2%	25,0%	28,0%	33,2%
	Sim	<b>95,5%</b>	26,5%	77,8%	75,0%	72,0%	66,8%
Acessórios Pessoais (Relógios)	Não	29,5%	45,6%	61,9%	6,3%	<b>76,0%</b>	43,3%
	Sim	<b>70,5%</b>	54,4%	38,1%	<b>93,8%</b>	24,0%	56,7%
Automóvel ou outro veículo motorizado	Não	84,1%	92,6%	92,1%	96,9%	100,0%	93,4%
	Sim	<b>15,9%</b>	7,4%	7,9%	3,1%	0,0%	6,6%
Brinquedos	Não	47,7%	<b>98,5%</b>	<b>93,7%</b>	56,3%	64,0%	74,4%
	Sim	52,3%	1,5%	6,3%	43,8%	36,0%	25,6%
Livros	Não	4,5%	17,6%	<b>68,3%</b>	31,3%	12,0%	28,7%
	Sim	<b>95,5%</b>	82,4%	31,7%	68,8%	<b>88,0%</b>	71,3%
Experiências / Vouchers	Não	31,8%	<b>97,1%</b>	25,4%	42,2%	<b>76,0%</b>	55,7%
	Sim	<b>68,2%</b>	2,9%	<b>74,6%</b>	57,8%	24,0%	44,3%
Filmes/Séries	Não	45,5%	86,8%	<b>98,4%</b>	85,9%	54,0%	77,2%
	Sim	54,5%	13,2%	1,6%	14,1%	46,0%	22,8%
Roupa	Não	25,0%	44,1%	<b>68,3%</b>	9,4%	88,0%	46,4%
	Sim	<b>75,0%</b>	55,9%	31,7%	<b>90,6%</b>	12,0%	53,6%
Tecnologia	Não	13,6%	<b>72,1%</b>	60,3%	9,4%	6,0%	35,3%
	Sim	<b>86,4%</b>	27,9%	39,7%	<b>90,6%</b>	<b>94,0%</b>	64,7%
Produtos para o Lar e Decoração	Não	38,6%	<b>91,2%</b>	<b>88,9%</b>	42,2%	78,0%	69,6%
	Sim	<b>61,4%</b>	8,8%	11,1%	57,8%	22,0%	30,4%
Refeições	Não	11,4%	<b>89,7%</b>	79,4%	85,9%	<b>94,0%</b>	75,4%
	Sim	<b>88,6%</b>	10,3%	20,6%	14,1%	6,0%	24,6%
Viagens	Não	6,8%	<b>85,3%</b>	15,9%	51,6%	48,0%	44,3%
	Sim	<b>93,2%</b>	14,7%	<b>84,1%</b>	48,4%	52,0%	55,7%
Supermercado	Não	15,9%	85,3%	<b>100,0%</b>	<b>93,8%</b>	<b>98,0%</b>	82,0%
	Sim	<b>84,1%</b>	<b>14,7%</b>	0,0%	6,3%	2,0%	18,0%
Outros	Não	97,7%	85,3%	96,8%	98,4%	100,0%	95,2%
	Sim	2,3%	14,7%	3,2%	1,6%	0,0%	4,8%
Sexo	Não	25,0%	63,2%	65,1%	64,1%	20,0%	50,5%
	Sim	<b>75,0%</b>	36,8%	34,9%	35,9%	<b>80,0%</b>	49,5%
Total		15%	24%	22%	22%	17%	100,0%

Nota: A negrito identificaram-se as principais características dos grupo.

**Análise discriminante:**

Começou-se por se introduzir as variáveis anteriormente discriminadas através do método *Enter* (introduzindo todas as vantagens em simultâneo e retirando variáveis manualmente à posteriori), utilizando a variável “Sexo” como variável de grupo, dado que se pretende obter quais os produtos que permitem diferenciar os compradores por sexo.

**Tabela 8:** Modelo de Análise Discriminante inicialmente obtido para os produtos comprados *Online*, usando como variável de grupo o “Sexo”

	Wilks' Lambda	F	df1	df2	Sig.
Bilhetes para Concertos/Espétaculos	.966	10.023	1	287	.002
Acessórios Pessoais (Relógios, Malas, Perfumes...)	.966	10.021	1	287	.002
Automóvel ou outro veículo motorizado	.955	13.532	1	287	.000
Brinquedos	.967	9.667	1	287	.002
Livros	.998	.634	1	287	.427
Experiências / Vouchers	1.000	.099	1	287	.753
Filmes/Séries	.971	8.594	1	287	.004
Roupa	.974	7.766	1	287	.006
Tecnologia	.842	53.965	1	287	.000
Produtos para o Lar e Decoração	.999	.155	1	287	.694
Refeições	.988	3.543	1	287	.061
Viagens	.996	1.051	1	287	.306
Supermercado	.991	2.611	1	287	.107
Outros	.999	.344	1	287	.558

**Tabela 9:** Matriz de estrutura da Análise Discriminante inicialmente obtido para os produtos comprados *Online*, usando como variável de grupo o “Sexo”

	Function
	1
Tecnologia	.631
Automóvel ou outro veículo motorizado	.316
Bilhetes para Concertos/Espétaculos	.272
Acessórios Pessoais (Relógios, Malas, Perfumes...)	-.272
Brinquedos	.267
Filmes/Séries	.252
Roupa	-.239
Refeições	.162
Supermercado	.139
Viagens	.088
Livros	.068
Outros	.050
Produtos para o Lar e Decoração	-.034
Experiências / Vouchers	-.027

Uma vez que a amostra foi recolhida com base num processo de amostragem não probabilístico, o que impossibilita calcular a precisão das estimativas obtidas a partir da amostra, optou-se por valorizar o valor absoluto valores da matriz de estrutura ( $> 0,200$ ) ao invés do método tradicional para escolha das variáveis a entrar para o modelo (p-value associado ao Lambda de Wilks). Os coeficientes da matriz de estrutura não variam com a introdução ou remoção de outras variáveis, pelo que foram removidas em simultâneo todas as variáveis cujo valor absolute da matriz de estrutura fosse inferior a 0.200. Ainda assim, os resultados obtidos com este método seriam os mesmos do que utilizando o método do p-value para um nível de significância de 95%. Utilizando este raciocínio, foram excluídos os produtos livros, experiências/vouchers, produtos para o lar e decoração, refeições, viagens, supermercado e outros (Tabela 10):

**Tabela 10:** Modelo final de Análise Discriminante obtido para os produtos comprados *Online*, usando como Variável de grupo “Sexo”

	Wilks' Lambda	F	df1	df2	Sig.
Bilhetes para Concertos/Espétaculos	.966	10.023	1	287	.002
Acessórios Pessoais (Relógios, Malas, Perfumes...)	.966	10.021	1	287	.002
Automóvel ou outro veículo motorizado	.955	13.532	1	287	.000
Brinquedos	.967	9.667	1	287	.002
Filmes/Séries	.971	8.594	1	287	.004
Roupa	.974	7.766	1	287	.005
Tecnologia	.842	53.965	1	287	.000

O valor da estatística Lambda de Wilks para a variável tecnologia parece intuir a possível existência de diferenças entre os dois grupos. De modo a confirmar esta indicação obtiveram-se os valores dos coeficientes da função de Fisher associados aos produtos comprados *online* (Tabela 11):

**Tabela 11:** Coeficientes dos produtos comprados *online* na função de Fisher por sexo

	Sexo	
	Feminino	Masculino
Bilhetes para Concertos/Espétaculos	5,426	6,110
Acessórios Pessoais (Relógios, Malas, Perfumes...)	3,146	1,992
Automóvel ou outro veículo motorizado	16,844	18,966
Brinquedos	1,572	1,960
Filmes/Séries	4,731	5,288
Roupa	3,659	2,953
Tecnologia	4,934	7,037
(Constant)	-26,351	-31,501

Deste modo as funções de Fisher resultantes da análise discriminante são da forma:

$$D_{Fem}(x) = 5,426.bilhetes + 3,146.acessorios\_pessoais + 16,844.auto - 1,572.brinquedos + 4,731.filmes + 3,659.roupa + 4,934.tecnologia - 26,351$$

$$D_{Masc}(x) = 6,110.bilhetes + 1,992.acessorios\_pessoais + 18,966.auto - 1,960.brinquedos + 5,288.filmes + 2,953.roupa + 7,037.tecnologia - 31,501$$

Assim, e considerando as diferenças entre os coeficientes de cada uma das funções de Fisher, parece existir alguma diferença entre os produtos comprados *online* por homens e por mulheres. Os homens compram mais tecnologia do que as mulheres, e estas por sua vez adquirem mais acessórios pessoais do que os homens.

Através do modelo apresentado, podemos ainda calcular a percentagem de indivíduos bem classificados em cada um dos grupos (homem ou mulher), sendo que, como é possível verificar, 76,0% das mulheres são bem classificadas, assim como 76,2% dos homens.

**Tabela 12:** Percentagem de homens e mulheres classificados corretamente, através do modelo obtido para os produtos comprados *online*

	Sexo	Predicted Group Membership		Total
		Feminino	Masculino	
Original	Feminino	76,0	24,0	100,0
%	Masculino	23,8	76,2	100,0

a. 76,1% das observações iniciais foram corretamente classificadas.

### Regressão logística:

De modo a determinar a função *logit* começou-se por construir um modelo considerando a variável “Sexo” como variável dependente (0 – mulher; 1 – homem) e as restantes variáveis como independentes, tendo-se obtido os resultados apresentados na Tabela 13.

**Tabela 13:** Modelo inicial de Regressão Logística obtido para o sexo em função dos produtos comprados *online*

	B	S.E.	$\left  \frac{S.E.}{\beta} \right $	Exp(B)
Bilhetes para Concertos / Espéculos	1,080	,363	,336	2,945
Acessórios Pessoais (Relógios, Malas, Perfumes...)	-1,152	,332	,288	,316
Automóvel ou outro veículo motorizado	2,807	,896	,319	16,564
Brinquedos	,464	,367	,791	1,591
Livros	-,137	,311	,27	,872
Experiências/ Vouchers	-,605	,360	,595	,546
Filmes/ Séries	,597	,372	,623	1,817
Roupa	-,610	,310	,508	,543
Tecnologia	2,287	,342	,15	9,841
Produtos para o Lar e Decoração	-,693	,370	,534	,500
Refeições	,509	,401	,788	1,664
Viagens	-,254	,359	,413	,776
Supermercado	,401	,437	,9	1,493
Outros	,894	,683	,764	2,446
Constant	-1,195	,199	,167	,303

Pelos motivos apresentados anteriormente (amostra recolhida com processo de amostragem não probabilístico), optou-se por valorizar o valor absoluto dos coeficientes e o respectivo desvio-padrão  $\left( \left| \frac{S.E.}{\beta} \right| \right)$  ao invés do método tradicional para escolha das variáveis a entrar para o modelo (p-value associado à estatística de Wald) – Ratner,2010. Neste procedimento não foram testadas as interações entre as variáveis. Utilizando este raciocínio, chegou-se ao modelo apresentado na Tabela 14, do qual é possível retirar as seguintes conclusões:

- Um indivíduo que compre bilhetes para concertos/espectáculos tem aproximadamente 2 vezes mais possibilidades de ser do sexo masculino do que do sexo feminino;
- Um indivíduo que compre acessórios pessoais (Relógios, Malas, Perfumes...) tem cerca de 4 vezes mais possibilidades de ser do sexo feminino do que do sexo masculino;
- Um indivíduo que compre um automóvel ou outro veículo motorizado tem aproximadamente 14 vezes mais possibilidades de ser do sexo masculino do que do sexo feminino;
- Um indivíduo que compre um produto de tecnologia tem cerca de 8 vezes mais possibilidades de ser do sexo masculino do que do sexo feminino.

**Tabela 14:** Modelo de Regressão Logística obtido para o sexo em função dos produtos comprados *online*

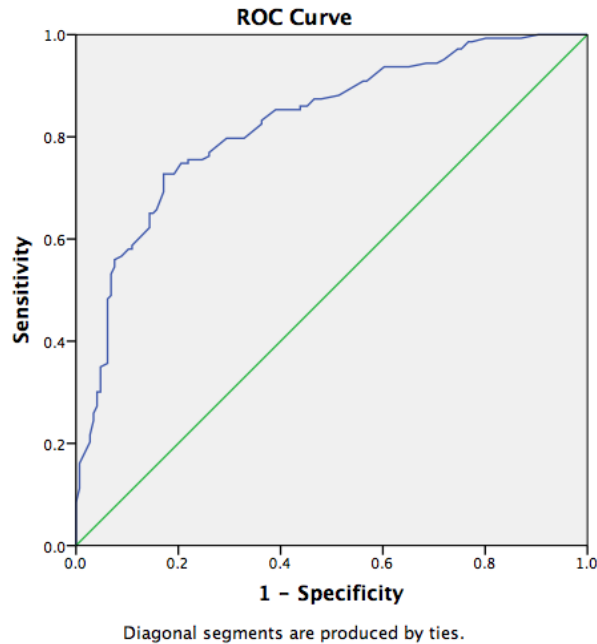
	B	S.E.	$\left  \frac{S.E.}{\beta} \right $	Exp(B)
Bilhetes para Concertos/ Espetáculos	.740	.275	.372	2.096
Acessórios pessoais(Relógios, Malas, Perfumes...)	-1.357	.299	.22	.257
Automóvel ou outro veículo motorizado	2.669	.825	.309	14.432
Tecnologia	2.070	.297	.143	7.924
Constant	-1.250	.186	.149	.286

Para estudar qualidade do ajustamento optou-se por utilizar os valores da sensibilidade e especificidade (Tabela 15). Assim, ao utilizarmos um ponto de corte de 0,50 obtemos um modelo com sensibilidade igual a 72,7% e uma especificidade igual a 1,1%, isto é, são identificados corretamente 72,7% dos homens e 80,8% das mulheres. No geral o modelo identifica corretamente 60,8% dos inquiridos o que representa uma boa qualidade de ajustamento. Podemos ainda verificar que o modelo encontrado tem uma boa capacidade discriminativa (AUC = 0,827).

**Tabela 15:** Valores previstos pelo modelo de regressão logística para o sexo em função dos produtos comprados *online*, usando o ponto de corte 0,50

Observed		Predicted		
		Sexo		Percentage Correct
		Feminino	Masculino	
Sexo	Feminino	118	28	8,8
	Masculino	39	104	72,7
Overall Percentage				76,8

**Gráfico 15:** Curva ROC associada ao modelo de regressão logística para o sexo em função dos produtos comprados *online*



Através dos valores  $\hat{\beta}$  apresentados na Tabela 8, podemos construir a função logística associada ao modelo obtido:

$$\begin{aligned} \text{logit}(\hat{\pi}(x)) &= \beta_0 + \beta_1 \cdot \text{bilhetes} - \beta_2 \cdot \text{acessorios\_pessoais} + \beta_3 \cdot \text{auto} - \beta_4 \cdot \text{tecnologia} \\ &= 0,740 \cdot \text{bilhetes} - 1,357 \cdot \text{acessorios\_pessoais} + 2,669 \cdot \text{auto} - 2,070 \cdot \text{tecnologia} - 1,250 \end{aligned}$$

### Comparação dos resultados obtidos

Um dos objetivos deste trabalho é estudar a coerência entre as variáveis e conclusões retiradas através da Análise de Correspondências Múltiplas, da função de Fisher resultante da Análise Discriminante e da Regressão Logística quando sujeitas ao mesmo problema. Neste sentido, e relativamente aos produtos comprados *online* por ambos os sexos, podemos concluir que quer através Análise de Correspondências Múltiplas (ACM) quer através da Função de Fisher resultante da Análise Discriminante (AD) conclui-se que poderão existir diferenças entre géneros no que diz respeito aos produtos comprados *online*.

Paralelamente, conclui-se ainda que no que diz respeito às variáveis identificadas como relevantes na Análise Discriminante (AD) e na Regressão Logística (RL), registou-se que todas as variáveis selecionadas para o modelo de RL também foram selecionadas para a análise discriminante.

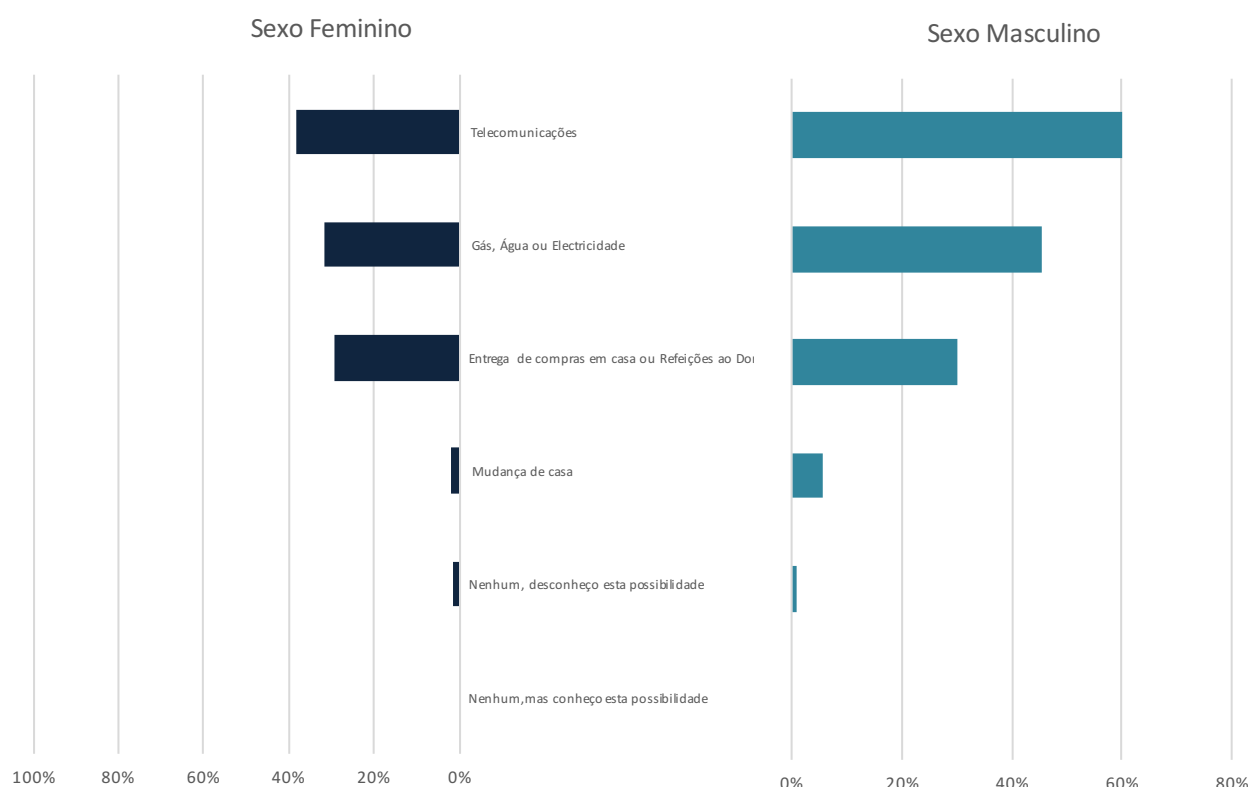
**Tabela 16:** Tabela resumo dos métodos utilizados na análise dos produtos comprados *Online*

Homens vs. Mulheres		
	Existem indícios de diferenças entre os grupos?	Variáveis diferenciadoras
Análise de Correspondências Múltiplas	Sim	<ul style="list-style-type: none"> <li>• Tecnologia</li> <li>• Bilhetes para Concertos / Espetáculos</li> <li>• Acessórios pessoais (Relógios, Malas, Perfumes...)</li> <li>• Experiências/Vouchers</li> <li>• Livros</li> <li>• Produtos para o Lar e Decoração</li> <li>• Viagens</li> <li>• Supermercado</li> </ul>
Análise Discriminante -> Função de Fisher	Sim	<ul style="list-style-type: none"> <li>• Tecnologia</li> <li>• Acessórios pessoais(Relógios, Malas, Perfumes...)</li> </ul>
Regressão Logística	Sim	<ul style="list-style-type: none"> <li>• Automóvel ou outro veículo motorizado</li> <li>• Tecnologia</li> <li>• Acessórios pessoais(Relógios, Malas, Perfumes...)</li> <li>• Bilhetes para Concertos / Espetáculos</li> </ul>



Interessou também analisar os hábitos de adesão a serviços *online* entre cada um dos géneros, sendo que à primeira vista sai que os homens aderem mais a serviços que as mulheres (74% dos homens já aderiram a pelo menos um serviço contra 64% das mulheres), essencialmente no que diz respeito a serviços de telecomunicações e a serviços de gás, à água ou electricidade, conforme é possível constatar no Gráfico 16. Note-se que a variável “Nenhum, mas conheço esta possibilidade” não foi escolhida por nenhum dos participantes no estudo.

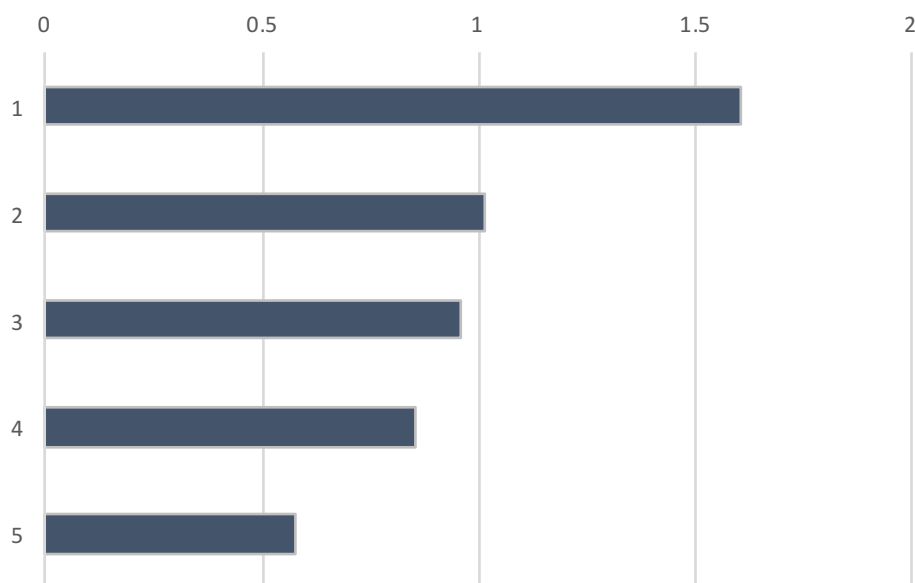
**Gráfico 16:** Serviços aderidos pelas mulheres e pelos homens



Seguindo procedimento análogo ao realizado para os produtos, estudou-se a variável sexo e a variável “A que serviços já aderiu *online*?”, por via a estudar a validade da **hipótese 3**: “Existem diferenças entre os serviços que homens e mulheres aderem *online*”.

### Análise de correspondências múltiplas:

A ACM valida uma solução com duas dimensões para representar as interações existentes entre as variáveis, as suas categorias e as respostas dos inquiridos (Gráfico 17). A percentagem de variância explicada pelas duas dimensões retidas é de 52,3%.

**Gráfico 17:** Representação da variância das dimensões associadas aos serviços aderidos *online***Tabela 17:** Medidas de discriminação associadas à ACM dos serviços aderidos *online*

	Dimension		Mean
	1	2	
Telecomunicações	.539	.040	.289
Entrega de compras em casa ou Refeições ao Domicilio	.256	.184	.220
Gás, Água ou Electricidade	.483	.099	.291
Mudança de casa	.240	.390	.315
Nenhum, desconheço esta possibilidade	.083	.298	.191
Sexo	.039	.002	.021
Active Total	1.602	1.012	1.307
% of Variance	32.036	20.240	26.138

Com base nos valores das medidas de discriminação e considerando as variáveis que mais contribuem dimensões, passaremos a denominar cada uma das dimensões anteriores do seguinte modo:

- Dimensão 1: **“Serviços domésticos”** (variáveis “Telecomunicações”, “Gás, Água ou Electricidade” e “Entrega de compras em casa ou Refeições ao domicilio” e “Mudança de casa”)
- Dimensão 2: **“Desconhecimento”** (variável “Nenhum, desconheço esta possibilidade”)

Através da Tabela 17, podemos ainda verificar que os valores das medidas de discriminação da variável sexo são próximos de zero para ambas as dimensões, o que indicia a não existência de diferenças entre géneros no que se refere à adesão *online* aos serviços.

**Tabela 18:** Posicionamento das variáveis mais relevantes para a ACM dos serviços aderidos *online*

	Quantificações negativas	Quantificações positivas
<u>Serviços domésticos</u> (Dimensão 1)	Telecomunicações_Sim Gás, Água ou Electricidade_Sim Entrega de compras em casa ou Refeições ao domicilio_Sim Mudança de casa_Sim	Telecomunicações_Não Gás, Água ou Electricidade_ Não Entrega de compras em casa ou Refeições ao domicilio_ Não Mudança de casa_ Não
<u>Desconhecimento</u> (Dimensão 3)	Nenhum, desconheço esta possibilidade_Não	Nenhum, desconheço esta possibilidade_Sim

Com base na Tabela 18, verifica-se que a dimensão **Serviços Domésticos** contempla todos os indivíduos que concordam relativamente à adesão *online* a serviços de Telecomunicações, Gás, Água ou Electricidade e Entrega de compras em casa ou Refeições ao domicílio e Mudança de casa. Já a dimensão **Desconhecimento** caracteriza-se pelos indivíduos que concordam não conhecer este tipo de adesão a serviços.

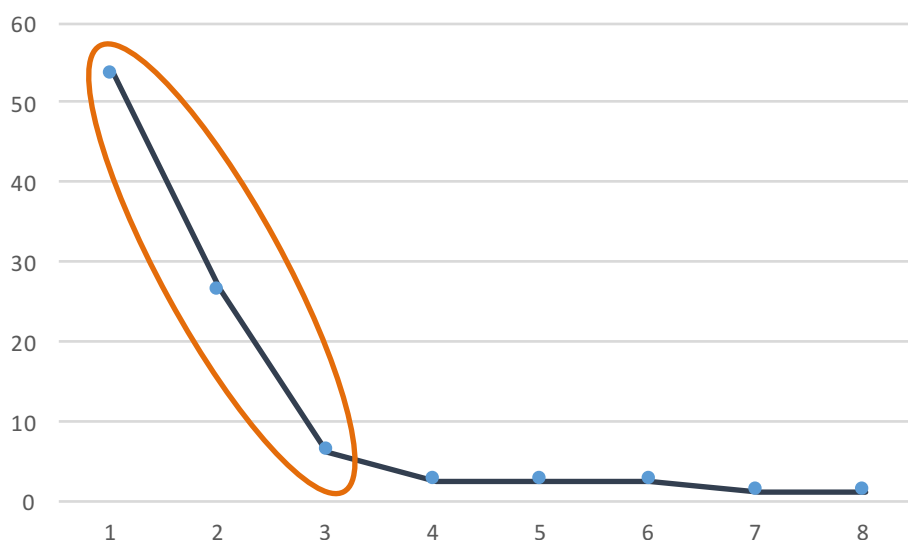
**Gráfico 18:** Representação do espaço dos serviços aderidos *online* - Dimensão 1 vs Dimensão 2

Da análise do Gráfico 18 é possível identificar 2 perfis, tendo em conta as variáveis que mais discriminam as respectivas dimensões, distintos de compradores:

1. Indivíduos que aderem *online* aos serviços mencionados;
2. Indivíduos que não aderem *online* a serviços;

Destaca-se ainda que as coordenadas dos centroides para os dois géneros, se encontram bastante próximas entre si (e da origem), o que, mais uma vez indica, a não existência de diferenças entre os dois grupos no que diz respeito aos serviços aderidos *online*.

De modo a validar a indicação dada pelo gráfico, realizou-se uma análise de clusters hierárquica dos scores dos indivíduos pelo modelo obtido de modo a determinar o número de grupos presentes.

**Gráfico 19:** Coeficientes de fusão da análise de Clusters associada aos serviços aderidos *online*

A análise de *clusters* hierárquica sugere com 3 grupos (Gráfico 19). Com base no agrupamento não hierárquico K-means, os grupos identificados caracterizam-se por (Tabela 19):

- O grupo 1 (95% dos compradores *online*) engloba os indivíduos que aderem *online* aos serviços relacionados com as despesas mensais da habitação.
- Já o 2º grupo (1% dos compradores *online*) distingue-se por não aderir *online* a qualquer tipo de serviços e nem conhecem essa possibilidade.
- O 3º grupo representa indivíduos que não se destacam na adesão a serviços (perfil médio).

**Tabela 19:** Caracterização dos grupos identificados, relativamente aos aos serviços aderidos *online*

	Categoria	Grupo 1	Grupo 2	Grupo 3	Total
Telecomunicações	Não	0.0%	<b>100.0%</b>	52.4%	50.9%
	Sim	<b>100.0%</b>	0.0%	47.6%	49.1%
Entrega de compras em casa ou Refeições ao Domicílio	Não	27.3%	<b>100.0%</b>	71.6%	70.2%
	Sim	<b>72.7%</b>	0.0%	28.4%	29.8%
Gás, Água ou Electricidade	Não	45.5%	<b>100.0%</b>	61.8%	61.6%
	Sim	54.5%	0.0%	38.2%	38.4%
Mudança de casa	Não	0.0%	100.0%	100.0%	96.2%
	Sim	<b>100.0%</b>	0.0%	0.0%	3.8%
Nenhum, desconheço esta possibilidade	Não	100.0%	0.0%	100.0%	99.0%
	Sim	0.0%	<b>100.0%</b>	0.0%	1.0%
Sexo	Feminino	27.3%	66.7%	51.3%	50.5%
	Masculino	<b>72.7%</b>	33.3%	48.7%	49.5%
Total		<b>95 %</b>	<b>1 %</b>	<b>4 %</b>	<b>100%</b>

Nota: A negrito identificaram-se as principais características dos grupo.

### Análise discriminante:

De entre as variáveis selecionadas com base no valor absoluto valores da matriz de estrutura, verifica-se pela estatística Lambda de Wilks que a mudança de casa não parece ser um serviço diferenciador entre os géneros, no que se refere à adesão de serviços *online* (Tabela 19).

**Tabela 20:** Modelo final de Análise Discriminante obtido para os serviços aderidos *online*, usando como variável de grupo "Sexo"

	Wilks' Lambda	F	df1	df2	Sig.
Telecomunicações	.953	14.300	1	287	.000
Gás, Água ou Electricidade	.979	6.024	1	287	.015
Mudança de casa	.991	2.476	1	287	.117

**Tabela 21:** Coeficientes da função de Fisher associada à Análise Discriminante os serviços aderidos *Online*, usando como variável de grupo "Sexo"

	Sexo	
	Feminino	Masculino
Telecomunicações	2.362	3.119
Gás, Água ou Electricidade	4.225	4.537
Mudança de casa	26.214	26.776
(Constant)	-18.482	-20.627

Com base na Tabela 21, as funções de Fisher resultantes da análise discriminante são da forma:

$$D_{Fem}(x) = 2,362 \cdot \text{telecomunicações} + 4,225 \cdot \text{gas\_agua\_electricidade} + 26,214 \cdot \text{mudanca\_casa} - 18,482$$

$$D_{Masc}(x) = 3,119 \cdot \text{telecomunicações} + 4,537 \cdot \text{gas\_agua\_electricidade} + 26,776 \cdot \text{mudanca\_casa} - 20,627$$

Assim e considerando as pequenas diferenças entre os coeficientes de cada uma das funções de Fisher, não existem indícios da existência diferenças relevantes entre homens e mulheres por serviços aderidos *online* por.

Com base no modelo apresentado, é possível verificar que 61,6% das mulheres são bem classificadas, assim como 69,1% dos homens, ou seja, 60,9% de todos os indivíduos foram classificados correctamente.

**Tabela 22:** Percentagem de homens e mulheres classificados corretamente, através do modelo obtido para os serviços aderidos *online*

	Sexo	Predicted Group Membership		Total
		Feminino	Masculino	
%	Feminino	61.6	38.4	100.0
	Masculino	39.9	60.1	100.0

a. 60.9% das observações iniciais foram corretamente classificadas.

### Regressão logística:

De modo a determinar a função *logit* começou-se por construir um modelo de regressão logística considerando a variável “Sexo” como variável dependente e os vários serviços disponibilizados como variáveis independentes. Foram introduzidos todos os serviços em simultâneo e foi feita uma seleção manual das variáveis utilizando o índice  $\left| \frac{S.E.}{\beta} \right|$ , sendo que quanto menores forem os valores deste rácio, melhor será a contribuição da variável para o modelo. Utilizando este raciocínio, chegou-se ao seguinte modelo apresentado na Tabela 23, do qual é possível verificar que aderir a serviços de telecomunicações ou de gás, água e electricidade é um factor potenciador para ser do sexo masculino.

**Tabela 23:** Modelo de Regressão Logística final obtido para o sexo em função dos serviços aderidos *online*

	B	S.E.	$\left  \frac{S.E.}{\beta} \right $	Exp(B)
Telecomunicações	.778	.258	.332	2.176
Gás, água, electricidade	.306	.266	.869	1.358
Constant	-1.604	.445	.277	.201

Ao utilizarmos um ponto de corte de 0,50 obtemos um modelo com sensibilidade igual a 60,1% e uma especificidade igual a 61,6%, isto é, são identificados corretamente 60,1% dos homens e 60,6% das mulheres, concluindo-se que o modelo se ajusta corretamente 60,9% dos inquiridos o que representa uma

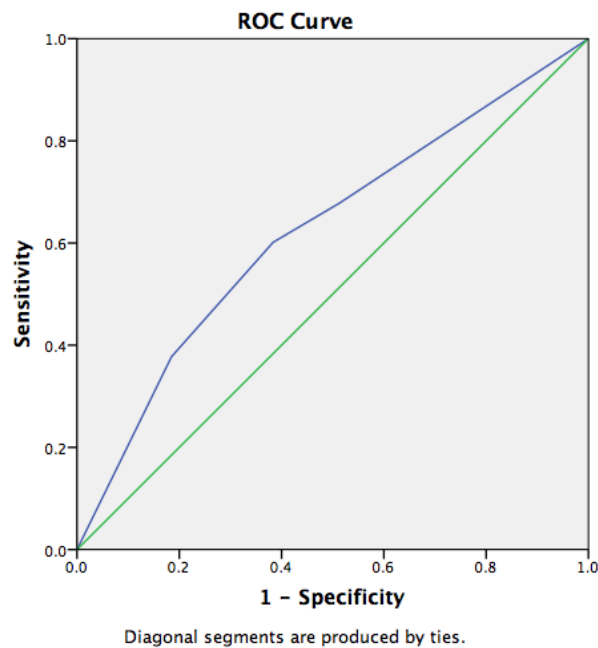
boa qualidade de ajustamento. O modelo encontrado tem ainda uma fraca capacidade discriminativa (AUC = 0,624).

**Tabela 24:** Valores previstos pelo modelo de regressão logística para o sexo em função dos serviços aderidos *online*, usando o ponto de corte 0,50

Observed		Predicted		
		Sexo		Percentage Correct
		Feminino	Masculino	
Sexo	Feminino	90	56	61.6
	Masculino	57	86	60.1
Overall Percentage				60.9

a. Ponto de corte = .500

**Gráfico 20:** Curva ROC associado ao Modelo de Regressão Logística obtido para os serviços aderidos Online, usando como variável suplementar "Sexo"



Através dos valores  $\hat{\beta}$ , podemos construir a função logística associada ao modelo obtido:

$$\begin{aligned} \text{logit}(\hat{\pi}(x)) &= \beta_0 + \beta_1 \cdot \text{telecom} + \beta_2 \cdot \text{gas\_agua\_electricidade} \\ &= 0,778 \cdot \text{telecomunica\c{c}oes} + 0,306 \cdot \text{gas\_agua\_electricidade} - 1,604 \end{aligned}$$



## Comparação dos resultados obtidos

Relativamente aos serviços aderidos *online* por ambos os sexos, foi possível concluir-se que através da ACM e da AD que não existem diferenças entre os géneros, porém, foram encontradas diferenças entre os grupos através da Regressão Logística.

De notar ainda que todas as variáveis escolhidas para o modelo de Regressão Logística também foram selecionadas para a Análise Discriminante.

**Tabela 25:** Tabela resumo dos métodos utilizados na análise dos serviços aderidos *Online*

	Homens vs. Mulheres	
	Existem indícios de diferenças entre os grupos?	Variáveis diferenciadoras
Análise de Correspondências Múltiplas	Não	
Análise Discriminante -> Função de Fisher	Não	
Regressão Logística	Sim	Telecomunicações

Outra das questões colocadas aos inquiridos foi acerca de para quem é que os utilizadores fazem compras *online*, se para si, se para outros, sendo que os resultados foram bastante similares entre homens e mulheres (Tabela 26).

**Tabela 26:** Hábitos de compra Online - para o próprio ou para outros

Faz compras <i>online</i> ...	% Homens	% Mulheres	Total
Para si	1,4 %	2,1%	1,7%
Para oferecer	37,8 %	36,9%	37,4%
Para si e para oferecer	60,8%	61,0%	60,9%

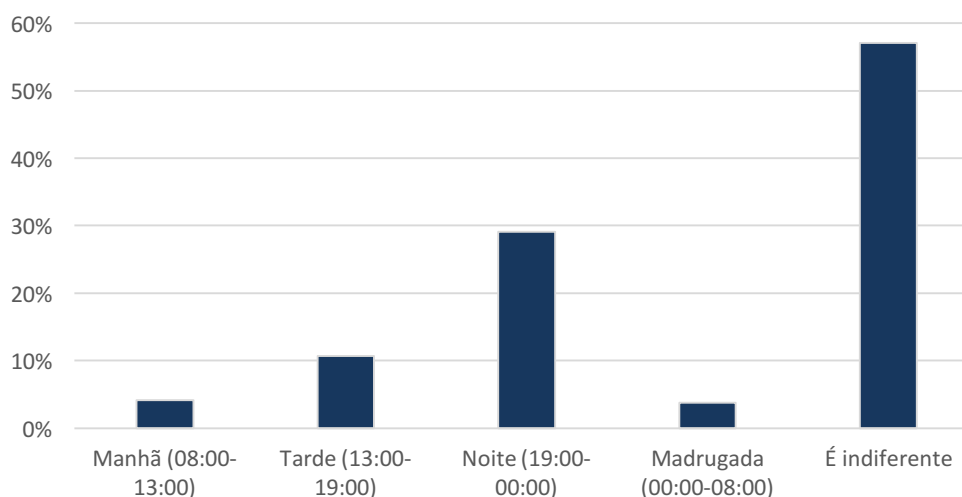
No que diz respeito à frequência de compras *online*, apenas 12,1% dos inquiridos que compram *online* refere fazer compras esporádicas (1 a 2 vezes por ano), sendo esta percentagem superior nas mulheres (15,8%), face aos homens (8,4%). Perto de 30% dos inquiridos indica que faz mais de 10 compras *online* por ano.

**Tabela 27:** Frequência das compras online por sexo

Faz compras <i>online</i> ...	% Homens	% Mulheres	Total
1 a 2 vezes	8,4%	15,8%	12,1%
3 a 5 vezes	25,9%	32,2%	29,1%
5 a 10 vezes	29,4%	28,8%	29,1%
Mais de 10 vezes	36,4%	23,3%	29,8%

Cerca de 80% dos inquiridos compradores *online* mostraram-se indiferentes ao dia da semana em que realizam a transação. Mais de metade (57,1%) dos inquiridos compradores dizem que não têm um período do dia em que é mais usual realizarem as compras. Já os restantes, mostram preferência pela noite ou madrugada (71,31%) sendo que o período do dia menos predileto para comprar *online* coincide com o período da manhã, onde apenas 4,15% dos inquiridos faz compras nesse período.

**Gráfico 21:** Período do dia predilecto para comprar *online*



Através das respostas dos indivíduos que responderam afirmativamente sobre o facto de comprar *online*, fica ainda a ideia de que as compras feitas na *web*, serão na sua maioria compras programadas, sendo que apenas 14,2% dos indivíduos disse fazer compras exclusivamente por impulso. Ainda assim 29% dos indivíduos que fazem compras programadas, admitem já terem feito compras por impulso.

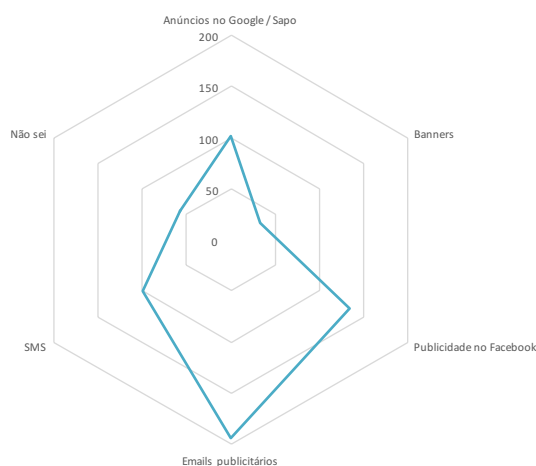
Como se pode verificar pela Tabela 28, entre os inquiridos compradores *online*, a percentagem de mulheres que compra por impulso é semelhante à dos homens (62,3% vs. 59,4%). Através do coeficiente de correlação de Phi (0,03), verifica-se que é quase inexistente a associação entre o sexo e o comprar por impulso, pelo que não existem diferenças entre géneros. De modo, neste conjunto de dados não se verifica a hipótese 4 “As inquiridas do sexo feminino têm maior tendência a comprar por impulso relativamente aos homens”.

**Tabela 28:** Programação prévia das compras por sexo

		Sexo		Total
		Feminino	Masculino	
Compras programadas	Não	91	85	176
	Sim	55	58	113
Total		146	143	289

No que concerne aos meios de publicidade *online*, nos quais os utilizadores tem maior consciência de terem sido influenciados comercialmente para fazerem compras, o *emailing* assume o protagonismo seguido da publicidade no Facebook (Gráfico 22).

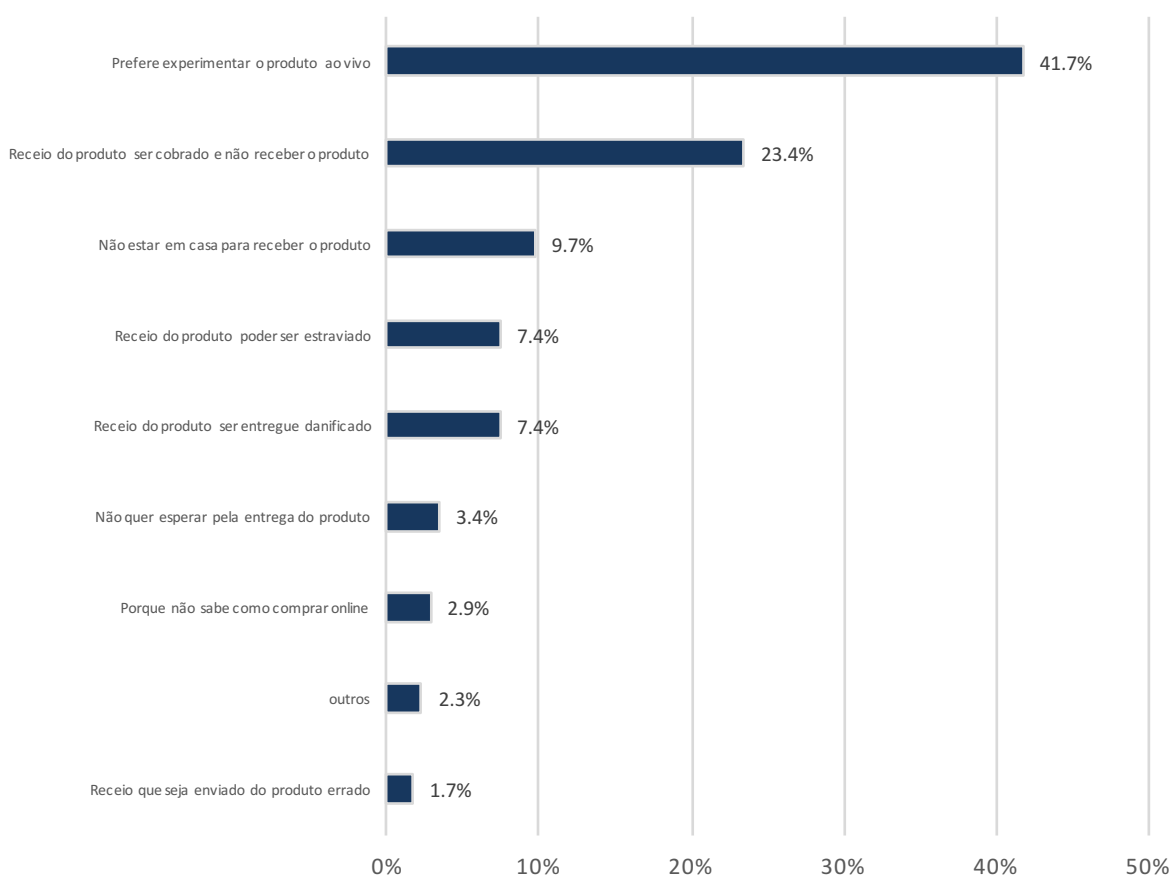
**Gráfico 22:** Meios de publicidade online que já motivaram compra



#### 4.4) Não Compradores *online*

Sendo este, um estudo sobre os hábitos de consumo de *e-commerce*, o principal grupo de estudo foi, naturalmente, os compradores *online*. Ainda assim, tentou-se perceber juntos dos não compradores *online* (28% dos inquiridos) quais as razões pelas quais estes não fazem compras através da *internet*. As principais razões apontadas, conforme é possível verificar no Gráfico 23, passam pela preferência por experimentar o produto ao vivo (41,7%) e o receio do produto ser cobrado e não ser entregue (23,4%). Verificou-se que menos de 2% dos inquiridos tem receio que seja enviado o produto errado e que 2,9% não sabe como comprar *online*.

**Gráfico 23:** Motivos pelos quais os inquiridos não compram *online*



Os restantes apontam razões como receio de extravio, produto ser entregue danificado, não querer esperar pela entrega do produto, dificuldades para estar em casa para receber o produto e outras razões.

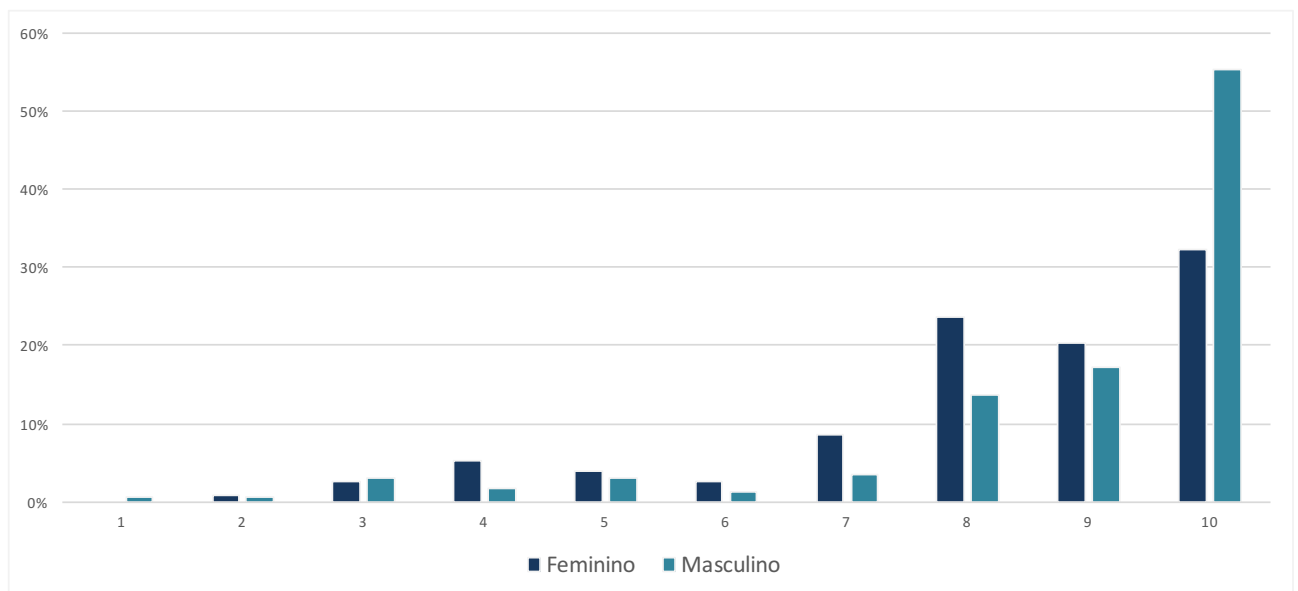
## 4.5) Segurança nas compras *online*

Um dos temas que surge inevitavelmente quando falamos de comércio eletrónico está relacionado com a segurança das transações online. Questões associadas à inserção de dados pessoais sensíveis, a possibilidade de ser alvo de ataques informáticos (*phishing*) e, o valor da compra ser elevado podem contribuir para algum receio no momento da transação. Porém, um maior número de escolhas de meios de pagamento disponíveis, a notoriedade e/ou a qualidade do site do vendedor podem servir como elementos securizantes que dão segurança durante a compra.

Este é então um tema muito relacionado com a perceção de cada utilizador e está diretamente ligado com as suas experiências pessoais ou de pessoas próximas.

Neste estudo, foi pedido aos inquiridos para indicarem o quão seguro consideram ser comprar online, respondendo estes numa escala ordinal, na qual 1 representava “Nada seguro” e 10 “Totalmente seguro”. Pelo menos metade dos inquiridos respondentes, afirmaram que consideravam este modo de compra seguro (mediana = 7). Os indivíduos do sexo masculino mostraram-se mais confiantes relativamente à segurança das transações ( $mediana_m = 10$ ) face aos indivíduos do sexo feminino ( $mediana_f = 9$ ).

**Gráfico 24:** Distribuição do grau de segurança nas compras *online* por sexo

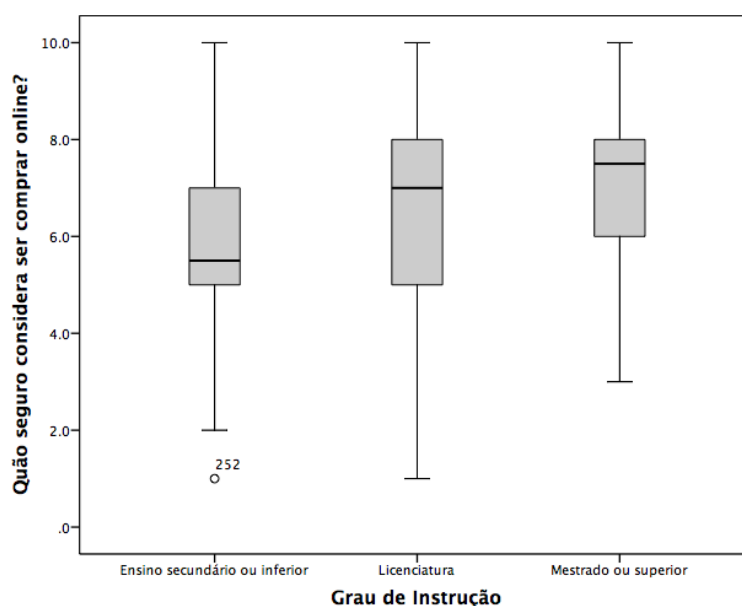


Os inquiridos entre os jovens e os indivíduos adultos foram os que indicaram sentir maior segurança nas compras *online* (mediana= 9,0).

**Gráfico 25:** Percentis do grau de segurança nas compras *online*, por idade categorizada

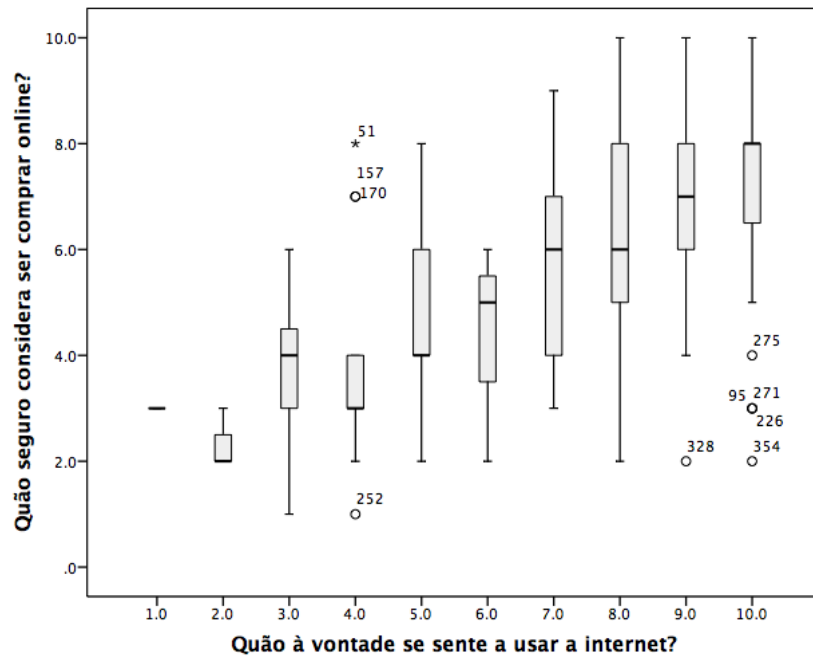
	Jovens-adultos	Adultos	Adultos maduros	Seniores
Mediana	9	9	8	8,5
Q1	8	8	7	8
Q3	10	10	9	9

Pretendeu-se ainda estudar o quão a percepção da segurança nas compras *online* varia de acordo com o nível de habilitações académicas. Em função da existência de poucos registos nas categorias extremas (5º ao 9º ano e Doutoramento), realizou-se um agrupamento dos dados do seguinte modo: as variáveis “5º ao 9º ano” e “10º ao 12º ano” foram aglutinadas numa só (nova denominação: “Ensino secundário e Inferior”) e a variável “Mestrado” foi aglomerada à variável “Doutoramento” (variável “Mestrado ou superior”), Deste modo, fica a percepção de que quanto maior for o grau académico maior a segurança sentida em comprar *online* (Gráfico 26). Analisando apenas a estatística descritiva associada a esta questão, existe indícios de uma tendência linear crescente do grau de segurança nas compras *online* em função do grau de instrução.

**Gráfico 26:** Caixa de bigodes “Grau de Instrução” vs. “Quão seguro considera ser comprar *online*?”

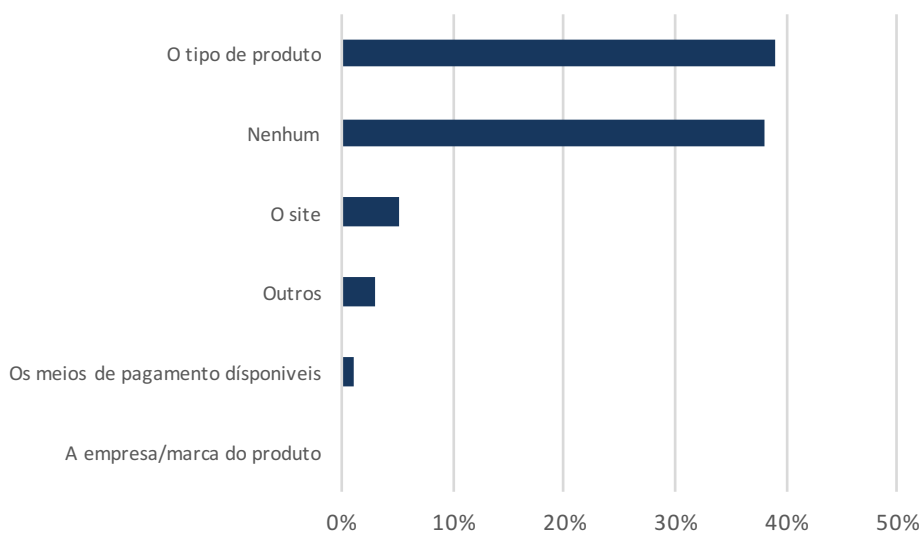
No seguimento do resultado anterior, surgiu um resultado semelhante: quanto mais alto é o nível de literacia persentido no uso da *internet*, mais alto será também o nível de perceção sobre a segurança nas compras *online*.

**Gráfico 27:** Caixa de bigodes do Grau de segurança nas compras *online* em função da auto-percepção de literacia no uso da *internet*



Para os inquiridos, os fatores mais importantes para a perceção de segurança na compra *online* são a o tipo de produto, nenhum e o site do retalhista (Gráfico 28).

**Gráfico 28:** Fatores de maior segurança nas compras *online*



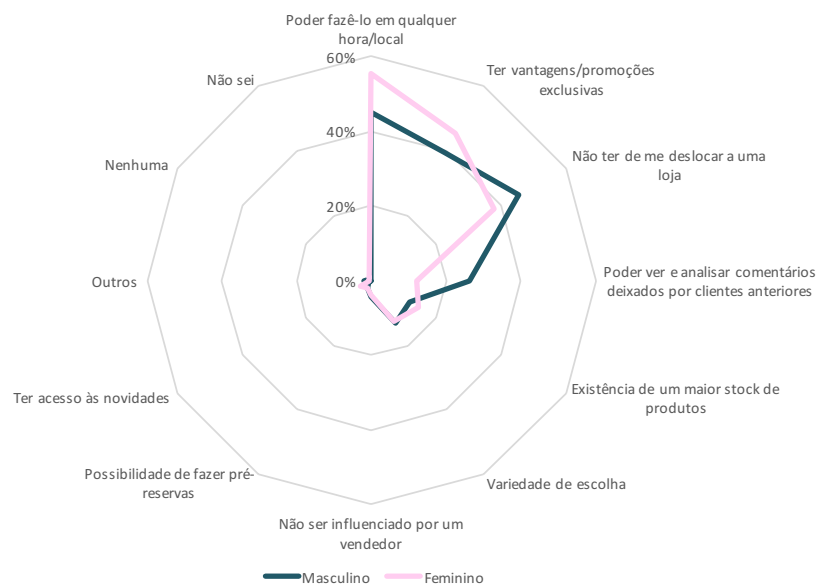
## 4.6) Preferências

Nesta secção serão abordadas as preferências associadas às compras *online* dos inquiridos (independentemente de já terem comprado *online* ou não) focando as principais vantagens e desvantagens que veem nas compras *online* e perceber como é estes acham que se poderá estimular mais o comércio eletrónico. Sempre que for relevante serão estudadas as diferenças entre compradores e não compradores e as diferenças de género.

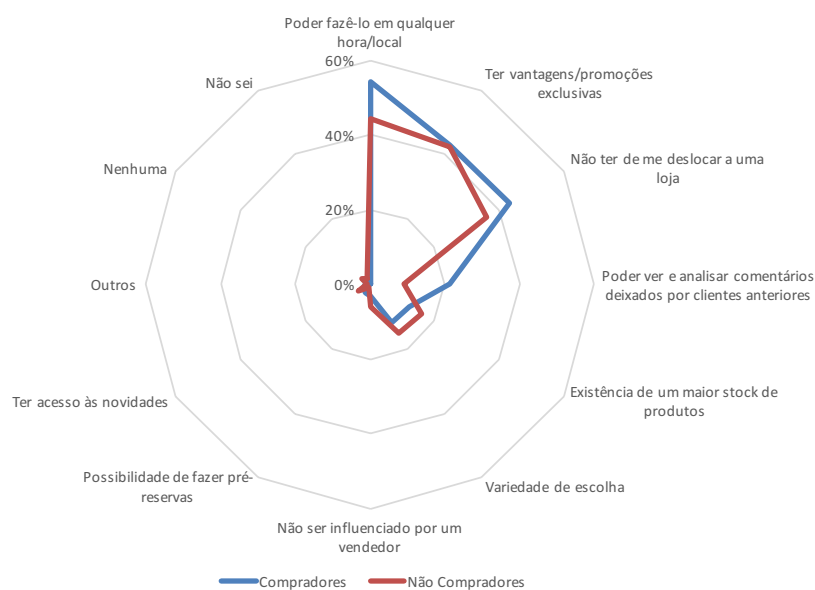
Pela análise do Gráfico 29, a generalidade dos utilizadores indicou que as principais vantagens ao comprar *online*, se prendiam com questões de comodidade, como sejam o poder fazê-lo a qualquer hora, em qualquer lugar e não se ter de deslocar a uma loja física (72%), e ainda existirem vantagens e promoções exclusivas (43%). De notar ainda que os indivíduos do sexo masculino valorizam bastante mais a possibilidade ler comentários deixados por clientes anteriores sobre um produto/serviço do que as mulheres (26,19% contra 12,07%).

### 4.6.1) Vantagens da compra *Online*

**Gráfico 29:** Vantagens associadas às compras *online* identificadas por homens e mulheres



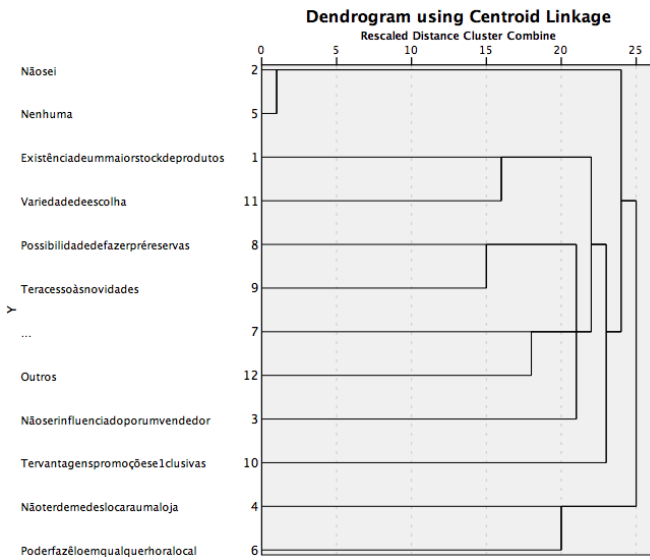


**Gráfico 30:** Vantagens associadas às compras *online* identificadas por compradores e não compradores

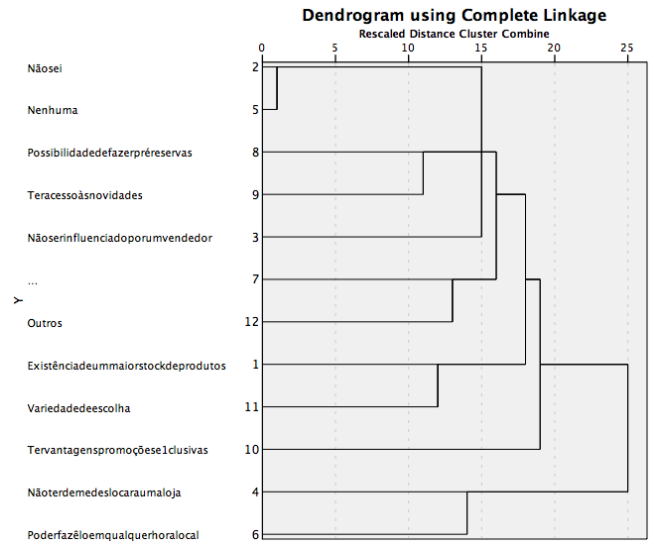
De modo a avaliar as diferenças entre as vantagens identificadas entre compradores e não compradores e entre géneros, do ponto de vista multivariado, serão usados os métodos: a Análise de Correspondências Múltiplas, a análise da função de Fisher associada à Análise Discriminante e a Regressão Logística. Porém, uma vez que existem variáveis com um número reduzido de observações, optou-se por se efetuar previamente uma Análise de *Clusters*, com carácter exploratório, de modo a identificar eventuais associações entre as variáveis por via a agrupar as variáveis com poucas observações.

Numa primeira fase traçou-se uma Análise de *Clusters* para a globalidade dos dados, usando vários métodos de classificação hierárquica: centroide, vizinho mais afastado e *Ward*, usando como medida de associação o coeficiente Phi, tendo-se obtido os dendogramas (Gráfico 31 a Gráfico 33).

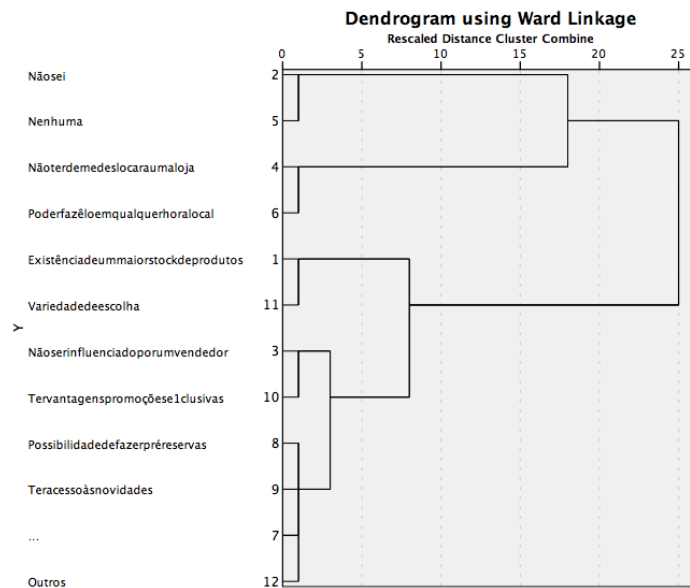
**Gráfico 31:** Dendograma das vantagens usando o método dos centroides



**Gráfico 32:** Dendograma das vantagens usando o método do vizinho mais afastado



**Gráfico 33:** Dendograma das vantagens usando o método de Ward

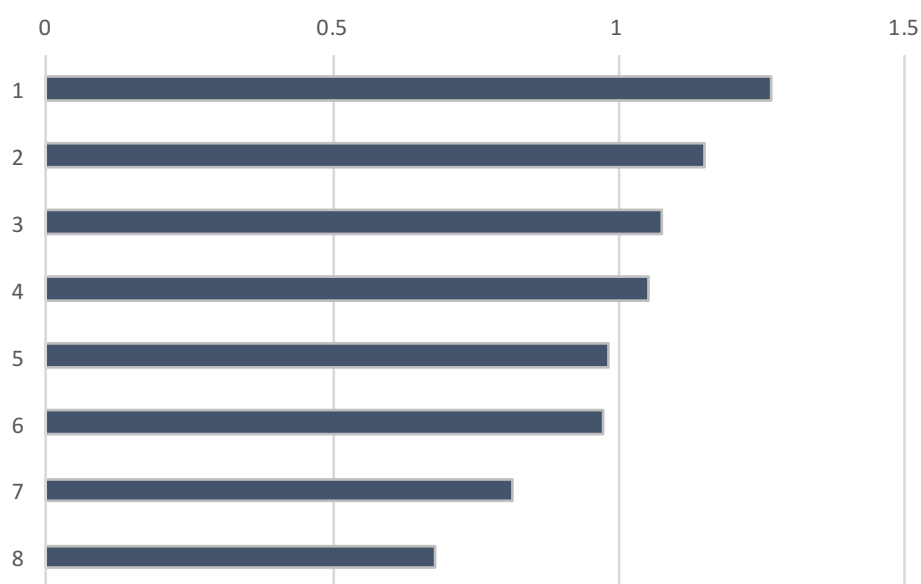


Através da análise dos dendogramas anteriores, optou-se por unificar as variáveis “Não ter que me deslocar a uma loja” e “Poder fazê-lo em qualquer hora/local” (que doravante será denominada por “Comodidade”), “possibilidade de fazer pré-reservas” e “Ter acesso às novidades” (que será denominada por “Novidades”) e as variáveis “nenhuma” e “não sei” (que será denominada por “Nenhuma / Não sei”).

### Análise de correspondências múltiplas:

Recorremos à ACM para realizar uma análise conjunta das vantagens identificadas na compra *online*. Simultaneamente verificaremos se existem diferenças entre as vantagens identificadas pelos compradores e pelos não compradores, bem como se existem diferenças entre géneros. Começou-se por definir o número máximo de dimensões que permitem representar o conjunto de dados, i.e., 8 (anexo 5). Em função dos resultados obtidos, para determinar o número de dimensões a reter optou-se por conjugar o método dos valores próprios (superiores a 1), o decaimento dos valores próprios (Gráfico 34) e a análise das medidas de discriminação obtidas (Tabela 29). Através destes métodos foram seleccionadas 4 dimensões, uma vez que a partir da 5ª dimensão, as variáveis relevantes para cada dimensão não adicionam nada de novo face às dimensões anteriores. Assim, as dimensões retidas representam 56,8% da variância total.

**Gráfico 34:** Representação da variância das dimensões associadas às vantagens da compra *online*



**Tabela 29:** Medidas de discriminação associadas à ACM das vantagens da compra *Online*

	Dimension				Mean
	1	2	3	4	
Não ser influenciado por um vendedor	.007	<b>.292</b>	.145	.001	.111
Poder ver e analisar comentários deixados por clientes anteriores	.100	.159	.018	<b>.526</b>	.201
Ter vantagens/promoções exclusivas	<b>.323</b>	.111	<b>.227</b>	.000	.165
Variedade de escolha	.158	.154	.026	<b>.417</b>	.189
Outros	.039	.055	.016	.019	.032
Nenhuma/Não sei	.188	<b>.215</b>	.034	.000	.109
Comodidade	<b>.403</b>	.155	.071	.000	.158
Novidades	.051	.008	<b>.539</b>	.088	.172
Costuma realizar compras Online? <sup>a</sup>	.001	.013	.002	.022	.009
Sexo <sup>a</sup>	.003	.014	.000	.016	.008
Active Total	1.269	1.149	1.077	1.051	1.137
% of Variance	15.868	14.369	13.457	13.134	14.207

a. Variável suplementar.

Com base nos valores das medidas de discriminação (Tabela 29), e considerando as variáveis que mais contribuem para a formação das dimensões, passaremos a denominar cada uma das dimensões anteriores do seguinte modo:

- Dimensão 1: **“Comodidade e exclusividade”** (variáveis “Comodidade” e “Ter vantagens/promoções exclusivas”)
- Dimensão 2: **“Influência de vendedores e nenhum/não sabe”** (variáveis: “Não ser influenciado por um vendedor” e “Nenhum/Não sei”)
- Dimensão 3: **“Compras rápidas ou de oportunidade”** (variáveis “Novidades” e “Ter vantagens/promoções exclusivas”)
- Dimensão 4: **“Comentários e variedade”** (variáveis: “Poder ver e analisar comentários deixados por clientes anteriores” e “Variedade de escolha”)

Note-se ainda que os valores das medidas de discriminação das variáveis “Costuma realizar compras *online*?” e “Sexo” são bastante próximas de zero para todas as variáveis, o que indicia a não existência de diferenças entre os elementos destes grupos.

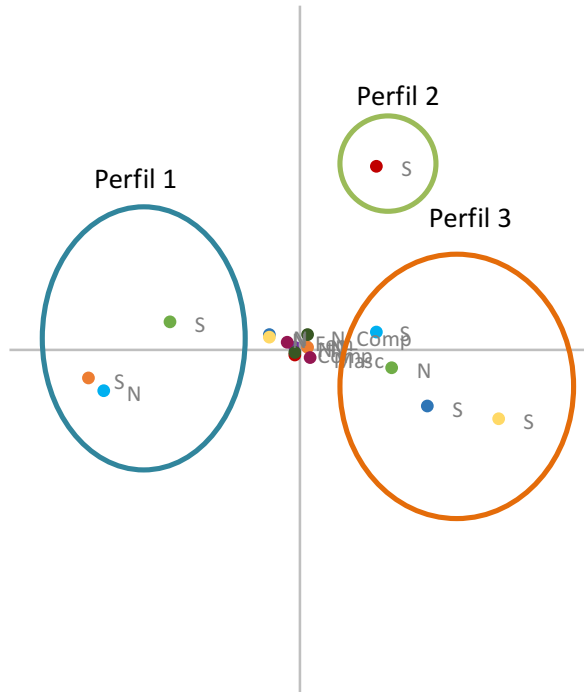
**Tabela 30:** Posicionamento das variáveis mais relevantes para a ACM das vantagens de compra *online*

	Quantificações negativas	Quantificações positivas
<u>Comodidade e Exclusividade</u> (Dimensão 1)	Comodidade_Não Ter vantagens e promoções exclusivas_Sim	Comodidade_Sim Ter vantagens e promoções exclusivas_Não
<u>Influência de vendedores e nenhum/não sabe</u> (Dimensão 2)	Não ser influenciado por um vendedor_Não Nenhum/Não sei_Sim	Não ser influenciado por um vendedor_Sim Nenhum/Não sei_Não
<u>Compras rápidas ou de oportunidade</u> (Dimensão 3)	Novidades_Não Ter_Vantagens/promoções_exclusivas_Sim	Novidades_Sim Ter_Vantagens/promoções_exclusivas_Não
<u>Comentários e variedade</u> (Dimensão 4)	Poder ver e analisar comentários deixados por clientes anteriores_Sim Variedade de escolha_Não	Poder ver e analisar comentários deixados por clientes anteriores_Não Variedade de escolha_Sim

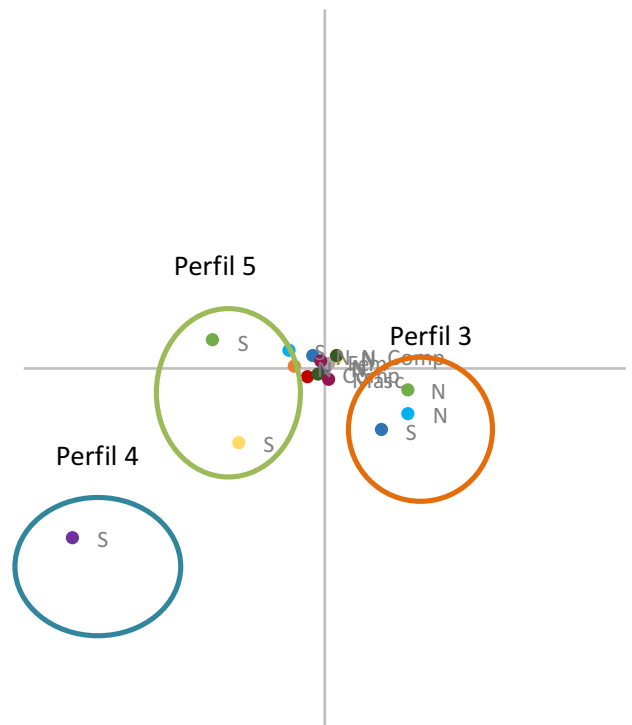
Com base na informação apresentada na Tabela 30, verifica-se que a dimensão **Comodidade e exclusividade** opõe os indivíduos que valorizam a Comodidade aos que preferem ter vantagens e promoções *online*. Já a dimensão **Influência de vendedores e nenhum/não sabe** caracteriza-se pelos indivíduos que são opostos em não ser influenciado por um vendedor e em não saber ou não encontrar nenhuma vantagem em comprar *online*. A dimensão **Compras rápidas ou de oportunidade** é a dimensão na qual os indivíduos interessados em ter acesso a pré-reservas e acesso as novidades se opõe aos indivíduos que preferem ter vantagens/promoções exclusivas. Finalmente, a dimensão **Feedback de clientes anteriores e variedade de escolha**, representa os indivíduos que discordam entre si entre ver e analisar comentários de clientes anteriores e a variedade de escolha.

Com base no valor das coordenadas dos centroides, representaram-se graficamente as dimensões mais relevantes quando comparadas duas a duas (Gráfico 35 e Gráfico 36):

**Gráfico 35:** Coordenadas dos centroides da ACM para as vantagens da compra *online*:  
Dimensão 1 vs Dimensão 2



**Gráfico 36:** Coordenadas dos centroides da ACM para as vantagens da compra *online*:  
Dimensão 2 vs Dimensão 3



- Não ser influenciado por um vendedor
- Poder ver e analisar comentários deixados por clientes anteriores
- Ter vantagens/promoções exclusivas
- Variedade de escolha
- Outros
- NenhumaNaosei
- Comodidade
- Novidades
- Costuma realizar compras Online?
- Sexo

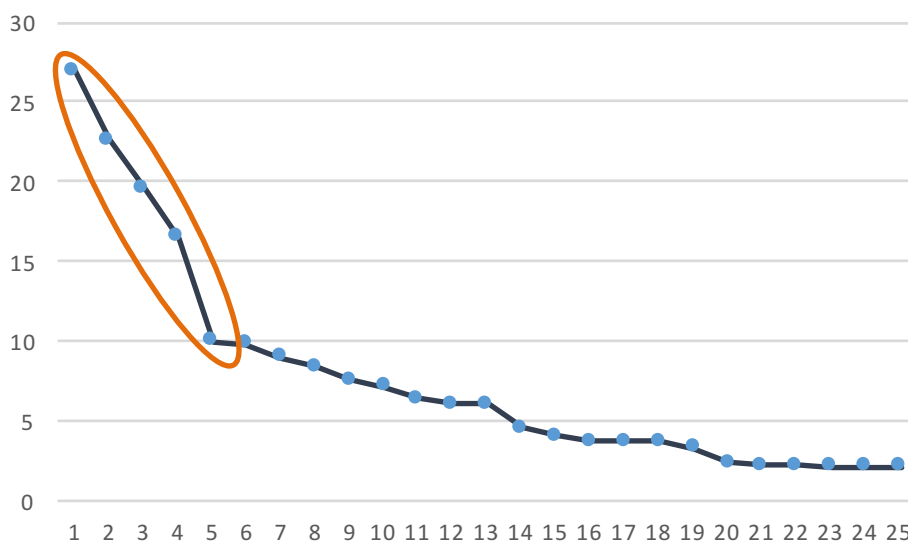
Da análise dos gráficos anteriores parecem existir, tendo em conta as variáveis que mais discriminam as respectivas dimensões, 5 perfis distintos de pessoas:

1. Indivíduos que destacam como vantagem as promoções exclusivas e as novidades e desvalorizam a comodidade;
2. Indivíduos que não gostam de ser influenciados por vendedores e outros;
3. Indivíduos que valorizam a opinião dos outros, a variedade disponível no *e-commerce* e a comodidade mas não valorizam as promoções exclusivas;
4. Indivíduos que valorizam outro tipo de vantagens que não as enunciadas no questionário;
5. Indivíduos que consideram como vantagem as promoções exclusivas e também a variedade na escolha.

Destaca-se ainda que as coordenadas dos centroides para os dois géneros, se encontram bastante próximas entre si (e da origem), o que, mais uma vez indica, a não existência de diferenças entre os grupos no que diz respeito aos às vantagens da compra *online*.

A Análise de Clusters hierárquica aplicada às coordenadas dos indivíduos nas dimensões retidas sugere o agrupamento dos indivíduos 5 grupos (Gráfico 37).

**Gráfico 37:** Coeficientes de fusão da análise de Clusters associada às vantagens da compra *online*



A partir da Análise de Clusters não hierárquica *K-means* é possível caracterizar os indivíduos de cada um dos grupos da seguinte forma (Tabela 31):

- Grupo 1 (16 % dos inquiridos): representa os indivíduos que valorizam a variedade de escolha, porém não valorizam vantagens/promoções exclusivas;
- Grupo 2 (36 % dos inquiridos): é caracterizado pelos indivíduos que não identificam nenhuma vantagem ou não sabem discriminar quais;
- Grupo 3 (25 % dos inquiridos): realça os indivíduos que não valorizam a variedade de escolha;
- Grupo 4 (15 % dos inquiridos): é marcado pelos indivíduos que valorizam a comodidade ao invés das vantagens/promoções exclusivas;
- Grupo 5 (8 % dos inquiridos): representa sobretudo elementos do sexo feminino que valorizam novidades por oposição à comodidade;

**Tabela 31:** Caracterização dos 5 grupos identificados, relativamente às vantagens da compra *online* identificadas pelos inquiridos.

	Categoria	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Total
Não ser influenciado por um vendedor	Não	100,0%	100,0%	100,0%	0,0%	100,0%	96,0%
	Sim	0,0%	0,0%	0,0%	<b>100,0%</b>	0,0%	4,0%
Poder ver e analisar comentários deixados por clientes anteriores	Não	84,3%	75,0%	80,8%	93,8%	87,5%	82,0%
	Sim	15,7%	25,0%	19,2%	6,3%	12,5%	18,0%
Ter vantagens/promoções exclusivas	Não	<b>72,5%</b>	25,0%	54,0%	<b>68,8%</b>	68,8%	57,3%
	Sim	27,5%	75,0%	46,0%	31,3%	31,3%	42,8%
Variedade de escolha	Não	5,9%	75,0%	<b>100,0%</b>	100,0%	87,5%	87,3%
	Sim	<b>94,1%</b>	25,0%	0,0%	0,0%	12,5%	12,8%
Outros	Não	92,2%	100,0%	100,0%	100,0%	100,0%	99,0%
	Sim	7,8%	0,0%	0,0%	0,0%	0,0%	1,0%
Nenhuma/Não sei	Não	100,0%	0,0%	100,0%	100,0%	100,0%	99,0%
	Sim	0,0%	<b>100,0%</b>	0,0%	0,0%	0,0%	1,0%
Comodidade	Não	17,6%	100,0%	28,8%	12,5%	<b>56,3%</b>	28,5%
	Sim	82,4%	0,0%	71,2%	<b>87,5%</b>	43,8%	71,5%
Novidades	Não	100,0%	100,0%	100,0%	93,8%	0,0%	95,8%
	Sim	0,0%	0,0%	0,0%	6,3%	<b>100,0%</b>	4,3%
Sexo	Não	54,9%	50,0%	58,1%	56,3%	<b>68,8%</b>	58,0%
	Sim	45,1%	50,0%	41,9%	43,8%	31,3%	42,0%
<b>Total</b>		13%	1%	78%	4%	4%	100%

Nota: A negrito identificaram-se as principais características dos grupo.



**Análise discriminante:**

Utilizando um procedimento análogo ao realizado anteriormente, começou-se por introduzir todas as variáveis no modelo, introduzindo todas as vantagens da compra *online* (usando facto de ser comprador ou não comprador *online* como variável de agrupamento) e fazendo uma seleção manual das variáveis utilizando o critério de seleção de variáveis com base nos valores da matriz de estrutura (> 0,200) – anexo 6).

**Tabela 32:** Modelo de Análise Discriminante para as vantagens na compra *Online*, usando como variável de agrupamento “Costuma realizar compras *online*”

	Wilks' Lambda	F	df1	df2	Sig.
Não ser influenciado por um vendedor	.995	2.129	1	398	.145
Poder ver e analisar comentários deixados por clientes anteriores	.979	8.552	1	398	.004
Variedade de escolha	.998	.906	1	398	.342
Nenhuma/Naosei	.996	1.550	1	398	.214
Comodidade	.995	1.943	1	398	.164
Novidades	.996	1.595	1	398	.207

Os valores da estatística Lambda de Wilks (próximos de 1 para quase todas as variáveis), parecem sugerir a não existência de diferenças entre os dois grupos (compradores e não compradores). De modo a verificar esta indicação estudou-se ainda as funções de Fisher associadas a este modelo (Tabela 33).

**Tabela 33:** Coeficientes da função de Fisher associada à Análise Discriminante para as vantagens da compra *Online*, usando como variável de agrupamento “Costuma realizar compras *online*”

	Costuma realizar compras <i>Online</i> ?	
	Não Compradores	Compradores
Não ser influenciado por um vendedor	28.316	27.602
Poder ver e analisar comentários deixados por clientes anteriores	8.022	8.857
Variedade de escolha	9.409	9.114
Nenhuma/Naosei	109.970	111.089
Comodidade	12.285	11.947
Novidades	30.402	29.670
(Constant)	-107.533	-107.185

Deste modo as funções de Fisher resultantes da análise discriminante são da forma:

$$D_{NComp}(x) = 28,316.nao\_ser\_influenciado\_por\_vendedor$$

$$+ 8,022.poder\_ver\_e\_analisar\_comentários\_deixados\_por\_clientes\_anteriores$$

$$+ 9,409.variedade\_de\_escolha + 109,970.nenhuma\_não\_sei + 12,285.comodidade$$

$$+ 30,402.novidades - 107,533$$

$$D_{Comp}(x) = 27,602.nao\_ser\_influenciado\_por\_vendedor$$

$$+ 8,857.poder\_ver\_e\_analisar\_comentários\_deixados\_por\_clientes\_anteriores$$

$$+ 9,114.variedade\_de\_escolha + 111,089.nenhuma\_não\_sei + 11,947.comodidade$$

$$+ 29,670.novidades - 107,185$$

Assim, e considerando as pequenas diferenças entre os coeficientes de cada uma das funções de Fisher, conclui-se que não existem diferenças relevantes entre compradores e não compradores relativamente às vantagens identificadas no *e-commerce*.

Através do modelo apresentado, podemos ainda calcular a percentagem de indivíduos bem classificados em cada um dos grupos (compradores ou não compradores), sendo que, como é possível verificar, 73,9% dos não compradores são bem classificados, porém apenas 42,2% dos compradores são bem classificados através deste modelo (Tabela 34). Verifica-se assim que o modelo apresenta deficiências na classificação dos compradores.

**Tabela 34:** Percentagem de compradores e não compradores classificados corretamente, através do modelo obtido para as vantagens da compra *online*

		Costuma realizar compras Online?	Predicted Group Membership		Total
			Não Compradores	Compradores	
Original	%	Não Compradores	73.9	26.1	100.0
		Compradores	57.8	42.2	100.0

a. 51.0% das observações iniciais foram corretamente classificadas.

Utilizando um raciocínio análogo, utilizando a variável sexo como variável de agrupamento, e usando o mesmo critério para escolha das variáveis, obtiveram-se os resultados apresentados na Tabela 35.

**Tabela 35:** Modelo de Análise Discriminante para as vantagens na compra *Online*, usando como variável de agrupamento "Sexo"

	Wilks' Lambda	F	df1	df2	Sig.
Poder ver e analisar comentários deixados por clientes anteriores	.967	13.545	1	398	.000
Ter vantagens/promoções exclusivas	.996	1.418	1	398	.234
Outros	.995	1.805	1	398	.180
Comodidade	.995	1.886	1	398	.170
Novidades	.997	1.153	1	398	.284

Tal como no exemplo anterior, os valores da estatística de Lambda de Wilks (próximos de 1 para todas as variáveis), parecem sugerir a não existência de diferenças entre os dois géneros. De modo a verificar esta indicação estudou-se ainda as funções de Fisher associadas a este modelo (Tabela 36).

**Tabela 36:** Coeficientes da função de Fisher associada à Análise Discriminante para as vantagens da compra *Online*, usando como variável de agrupamento "Sexo"

	Sexo	
	Feminino	Masculino
Poder ver e analisar comentários deixados por clientes anteriores	8.910	9.880
Ter vantagens/promoções exclusivas	8.305	8.146
Outros	103.045	104.364
Comodidade	9.869	9.478
Novidades	30.423	29.825
(Constant)	-88.053	-89.025

Assim, as funções de Fisher resultantes da análise discriminante são da forma:

$$D_{Fem}(x) = 8,910.poder\_ver\_e\_analisar\_comentários\_deixados\_por\_clientes\_anteriores + 8,305.ter\_vantagens\_e\_promocoes\_exclusivas + 103,045.outros + 9,869comodidade + 30,423.novidades - 88,053$$

$$D_{Masc}(x) = 9,880.poder\_ver\_e\_analisar\_comentários\_deixados\_por\_clientes\_anteriores + 8,146.ter\_vantagens\_e\_promocoes\_exclusivas + 104,364.outros + 9,478comodidade + 29,825.novidades - 89,025$$

Assim e considerando as pequenas diferenças entre os coeficientes de cada uma das funções de Fisher temos que não existem indícios de diferenças relevantes entre as vantagens identificadas pelos dois géneros, confirmando as orientações da estatística Lambda de Wilks.

Calculou-se, ainda, através da classificação efetuada a percentagem de indivíduos bem classificados em cada um dos grupos (homem ou mulher), sendo que 67,2% dos elementos do sexo feminino são bem classificados, porém apenas 51,8% dos homens são bem classificados através deste modelo (**Tabela 37**). Deste modo, constata-se que este modelo apresenta deficiências na classificação correcta dos homens.

**Tabela 37:** Percentagem de homens e mulheres classificados corretamente, através do modelo obtido às vantagens das compras *online*

		Sexo	Predicted Group Membership		Total
			Feminino	Masculino	
Original	%	Feminino	67.2	32.8	100.0
		Masculino	48.2	51.8	100.0

a. 60.8% das observações iniciais foram corretamente classificadas.

### Regressão logística:

Para construir um modelo de regressão logística para a variável “Sexo” considerando as vantagens na compra *online* como variáveis independentes, introduziram-se todas as vantagens em simultâneo e foi feita uma seleção manual das variáveis utilizando o índice  $\left| \frac{S.E.}{\beta} \right|$ , sendo que quanto menores forem os valores deste rácio, maior será a contribuição da variável para o modelo. Utilizando este raciocínio, chegou-se ao seguinte modelo apresentado na Tabela 38, do qual é possível verificar que os indivíduos que gostam de saber a opinião dos outros sobre os produtos a adquirir ou sobre as lojas onde adquirir é um factor potenciador para ser comprador *online*.

**Tabela 38:** Modelo de Regressão Logística para os compradores online em função das vantagens na compra online

	B	S.E.	$\left  \frac{S.E.}{\beta} \right $	Exp(B)
Poder ver e analisar comentários deixados por clientes anteriores	1.023	.362	.354	2.781
Comodidade	-.431	.263	.610	.650
Novidades	-.691	.520	.753	.501
Constant	1.255	.864	.688	3.507

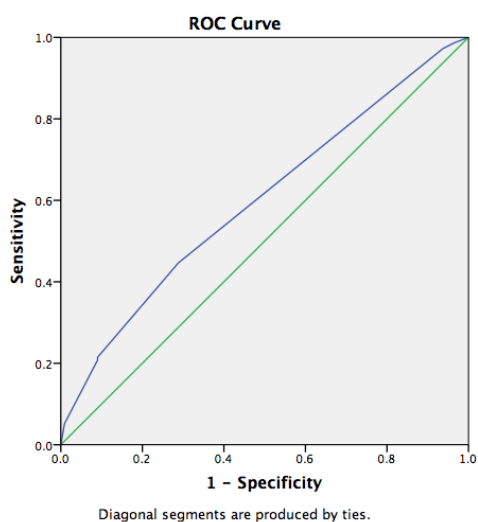
Para um ponto de corte de 0,70 obtemos um modelo com sensibilidade igual a 44,6% e uma especificidade igual a 71,2%, isto é, são identificados corretamente 44,6% dos compradores e 71,2% dos não compradores, concluindo que o modelo ajusta corretamente 52,0% dos inquiridos o que representa uma qualidade razoável de ajustamento. À semelhança do verificado com a Análise discriminantes, também este modelo apresenta deficiências na classificação correcta dos compradores. O modelo encontrado tem uma fraca capacidade discriminativa (AUC = 0,599).

**Tabela 39:** Valores previstos pelo modelo de regressão logística para os compradores em função das vantagens associadas à compra *online*, usando o ponto de corte 0,70

Observed		Predicted		
		Costuma realizar compras Online?		Percentage Correct
		Não Compradores	Compradores	
Costuma realizar compras Online?	Não Compradores	79	32	71.2
	Compradores	160	129	44.6
Overall Percentage				52.0

a. Ponto de corte = .700

**Gráfico 38:** Curva ROC associado ao Modelo de Regressão Logística obtido para as vantagens na compra Online, usando como variável suplementar “Costuma realizar compras online”



Através dos valores  $\hat{\beta}$ , podemos construir a função logística associada ao modelo obtido:

$$\text{logit}(\hat{\pi}(x)) = \beta_0 + \beta_1 \cdot \text{poder\_ver\_e\_analizar\_comentarios\_deixados\_por\_clientes\_anteriores} + \beta_2 \cdot \text{comodidade} + \beta_3 \cdot \text{novidades}$$

$$= 1,023.poder\_ver\_e\_analisar\_comentarios\_deixados\_por\_clientes\_anteriores - 0,431.comodidade + 0,691.novidades + 1,255$$

Repetindo o mesmo raciocínio, agora para modelar o sexo em função das vantagens identificadas na compra *online*, tanto as outras vantagens não discriminadas no questionário como o poder ter acesso aos comentários deixados pelos clientes anteriores são factores potenciadores para ser do sexo masculino (Tabela 40).

**Tabela 40:** Modelo de Regressão Logística para o sexo em função das vantagens

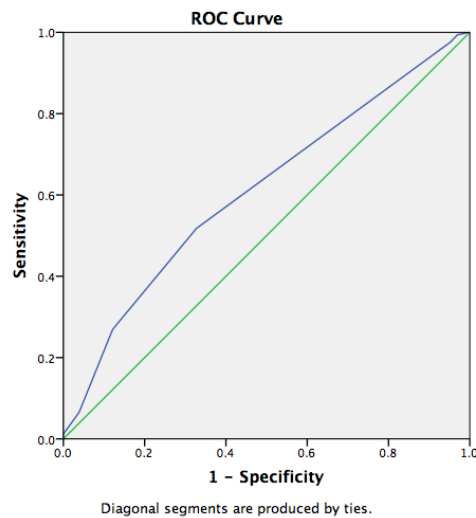
	B	S.E.	$\frac{ S.E. \beta}{\beta}$	Exp(B)
Poder ver e analisar comentários deixados por clientes anteriores	.961	.269	0,280	2.615
Outros	1.401	1.174	0,838	4.061
Comodidade	-.381	.230	-0,604	.683
Novidades	-.620	.558	-0,900	.538
Constant	-1.580	1.436	-0,909	.206

Ao utilizarmos um ponto de corte de 0,50 obtemos um modelo com sensibilidade igual a 27,4% e uma especificidade igual a 87,5%, isto é, são identificados corretamente 27,4% dos indivíduos do sexo masculino e 87,5% dos indivíduos do sexo feminino, concluindo que o modelo ajusta corretamente 62,3% da totalidade dos inquiridos. O modelo apresenta assim uma fraca capacidade em classificar correctamente os indivíduos do sexo masculino. Porém, o modelo encontrado tem uma fraca capacidade discriminativa (AUC = 0,610).

**Tabela 41:** Valores previstos pelo modelo de regressão logística para o sexo em função das vantagens associadas às compras *online*, usando o ponto de corte 0,50

Observed		Predicted		
		Sexo		Percentage
		Feminino	Masculino	Correct
Sexo	Feminino	203	29	87.5
	Masculino	122	46	27.4
Overall Percentage				62.3

a. Ponto de corte = .500

**Gráfico 39:** Curva ROC associado ao Modelo de Regressão Logística para o sexo em função das vantagens na compra *online*

Através dos valores  $\hat{\beta}$ , podemos construir a função logística associada ao modelo obtido:

$$\text{logit}(\hat{\pi}(x)) = \beta_0 + \beta_1 \cdot \text{poder\_ver\_e\_analisar\_comentarios\_deixados\_por\_clientes\_anteriores} + \beta_2 \cdot \text{outros} + \beta_3 \cdot \text{comodidade} + \beta_4 \cdot \text{novidades}$$

$$= 0,961 \cdot \text{poder\_ver\_e\_analisar\_comentarios\_deixados\_por\_clientes\_anteriores} + 1,401 \cdot \text{outros} - 0,381 \cdot \text{comodidade} - 0,620 \cdot \text{novidades} - 1,580$$

Por via a comparar as variáveis e conclusões retiradas através da Análise de Correspondências Múltiplas (ACM), da função de Fisher resultante da Análise Discriminante (AD) e da Regressão Logística (RL), quando sujeitas ao mesmo problema, a tabela resumo das vantagens da compra *online* identificadas por compradores e não compradores e entre géneros entre os vários modelos utilizados (Tabela 42) permite concluir que, quer através ACM quer através da função de Fisher resultante da AD, não existem diferenças entre compradores e não compradores e nem entre géneros. Paralelamente, conclui-se ainda que no que diz respeito às variáveis identificadas como relevantes na AD e na RL:

- As variáveis escolhidas na AD para os grupos construídos por homens/mulheres e compradores/não compradores, foram as mesmas;
- Todas as variáveis selecionadas para o modelo de RL para os compradores e não compradores são também relevantes para o modelo de RL para o sexo;
- Os modelos de AD e RL identificaram variáveis distintas no estudo para os grupos de compradores/não compradores e homens/mulheres;

**Tabela 42:** Tabela resumo dos métodos utilizados na análise das vantagens em comprar *Online*

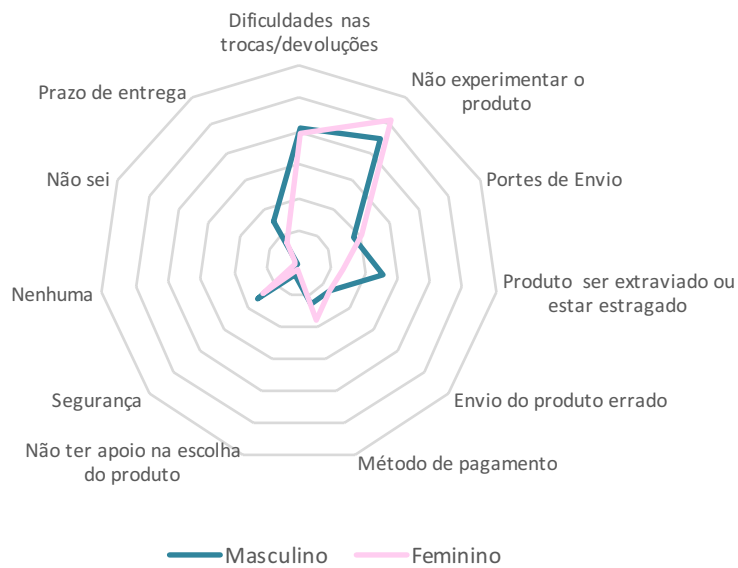
	Compradores vs. Não Compradores		Homens vs. Mulheres	
	Existem indícios de diferenças entre os grupos?	Variáveis diferenciadoras	Existem indícios de diferenças entre os grupos?	Variáveis diferenciadoras
Análise de Correspondências Múltiplas	Não		Não	
Análise Discriminante -> Função de Fisher	Não		Não	
Regressão Logística	Sim	<ul style="list-style-type: none"> <li>• Poder ver e analisar comentários deixados por clientes anteriores</li> </ul>	Sim	<ul style="list-style-type: none"> <li>• Poder ver e analisar comentários deixados por clientes anteriores</li> <li>• Outros</li> </ul>



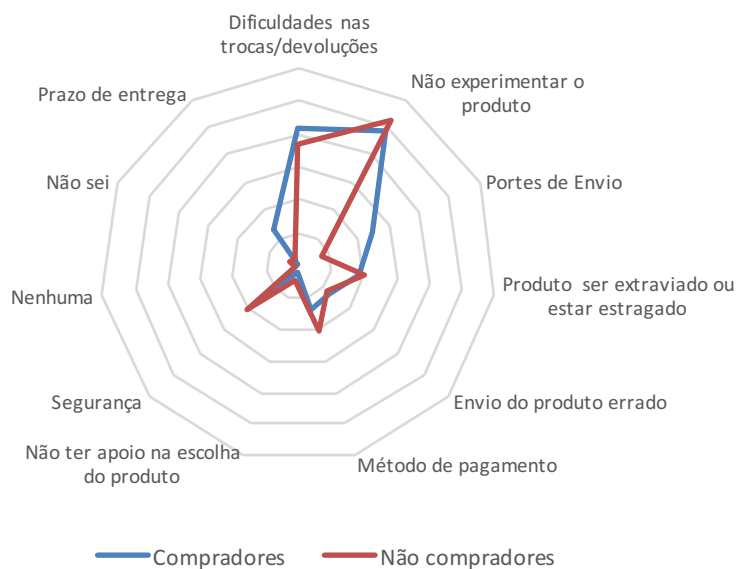
### 4.6.2) Desvantagens da compra *Online*

Segue-se agora a análise das desvantagens associadas ao *e-commerce* que foram identificadas pelos inquiridos. A desvantagem mais identificada pela totalidade dos inquiridos é a não experimentação do produto ao vivo, sendo esta mais expressiva nas mulheres (52,2% vs. 45,2% dos inquiridos do sexo masculino). Curioso notar que esta desvantagem é quase tão apontada pelos compradores (48,1%) como pelos não compradores (52,3%; Gráfico 41).

**Gráfico 40:** Desvantagens associadas às compras *Online* identificadas por homens e mulheres



**Gráfico 41:** Desvantagens associadas às compras *Online* identificadas por Compradores e Não Compradores



Os inquiridos de ambos os sexos (41,1% dos homens e 39,7% das mulheres) mostram também uma grande preocupação com eventuais dificuldades nas trocas e devoluções (Gráfico 38). Esta preocupação é superior entre os compradores (41,5%) do que nos não compradores (36,9%). Os homens preocupam-se mais (25,6%) com o extravio/produto ser entregue danificado do que as mulheres (13,8%).

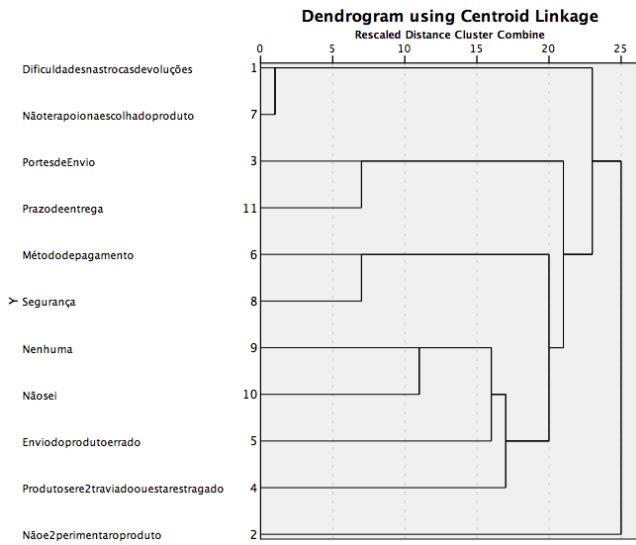
Estas diferenças apresentadas parecem sugerir uma relação entre a tipologia de produto comprado por cada um dos sexos (homens compram na sua maioria tecnologia, enquanto as mulheres optam pela moda e acessórios pessoais) e o tipo de preocupações que estes têm no ato da compra *online*.

De referir ainda, a existência de diferenças de género nas respostas dadas pelos inquiridos: as mulheres preocupam-se mais com os métodos de pagamento disponíveis do que os homens (18,1% vs. 12,5%), porém os homens revelaram maiores preocupações com o prazo de entrega (14,9% dos inquiridos do sexo masculino vs. 7,3% no sexo feminino).

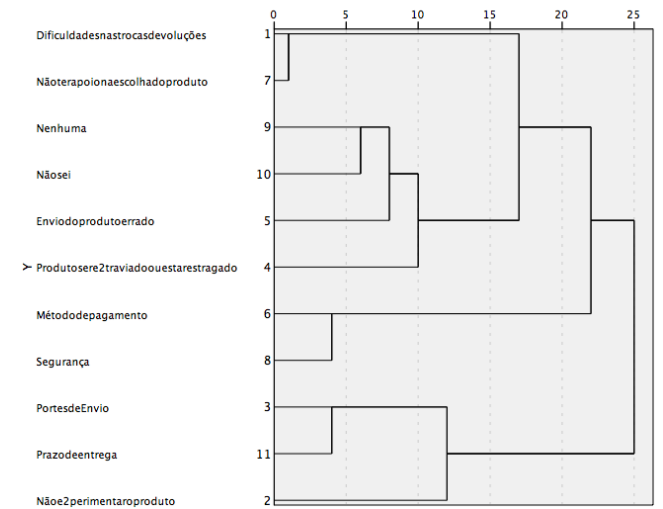
De notar, por fim, que os compradores *online* se preocupam bastante mais (24,9% vs. 7,2% dos não compradores) com a existência de portes de envio assim como com o prazo de entrega (13,5% vs. 2,7%). Portanto, denota-se nos compradores mais preocupações com aspetos práticos que o registado entre os não compradores. Pelo contrário, os não compradores têm preocupações superiores aos compradores com questões mais ligadas à segurança (20,7% dos inquiridos não compradores vs. 13,8% dos compradores) e com o método de pagamento (20,7% dos não compradores vs. 12,8% dos compradores).

De modo a averiguar as diferenças relativas às desvantagens de comprar *online* identificadas por compradores e não compradores e entre géneros, repetiu-se o processo de comparação de métodos realizado anteriormente. À semelhança do que foi feito na análise das vantagens, começou-se por se efetuar uma *Análise de Clusters*, com carácter exploratório, de modo a identificar eventuais associações entre as variáveis por via a agrupar variáveis com poucas observações.

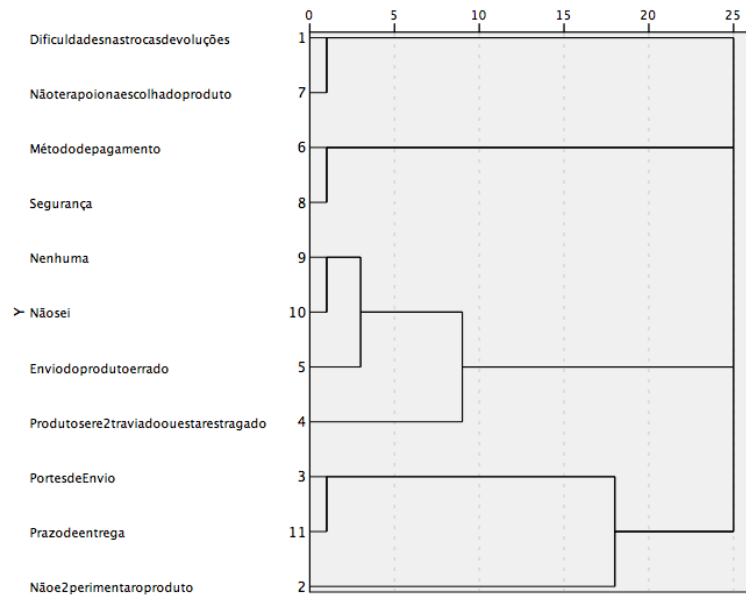
**Gráfico 42:** Dendrograma das desvantagens usando o método dos centroides



**Gráfico 43:** Dendrograma das desvantagens usando o método do vizinho mais afastado

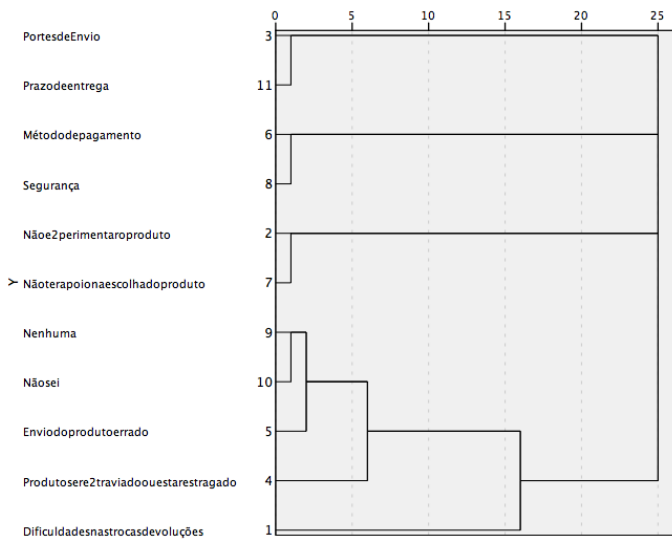


**Gráfico 44:** Dendrograma das desvantagens usando o método de Ward

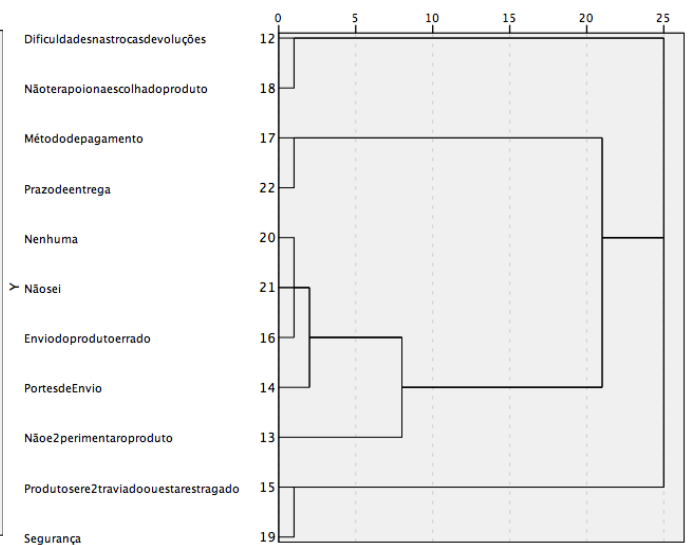


Através da análise dos dendogramas (Gráfico 42 a Gráfico 44), optou-se por unificar as variáveis “Dificuldades nas trocas e devoluções” e “Não ter apoio na escolha do produto” (que doravante será denominada por “Falta de Apoio no pré e pós venda”) e as variáveis “Nenhuma” e “Não sei” (que será denominada por “Nenhuma / Não sei”). A mesma unificação foi considerada com base na Análise de *Clusters* separada por sexo e por compradores e não compradores (Gráfico 45 a Gráfico 48).

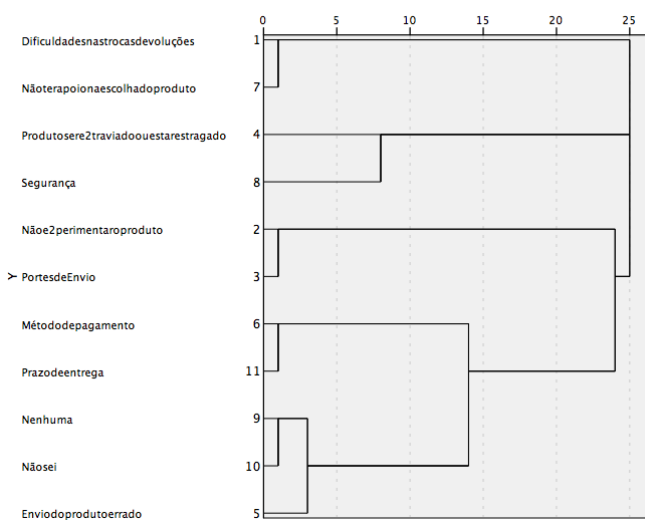
**Gráfico 45:** Dendograma das desvantagens usando o método de Ward para os inquiridos do sexo Feminino



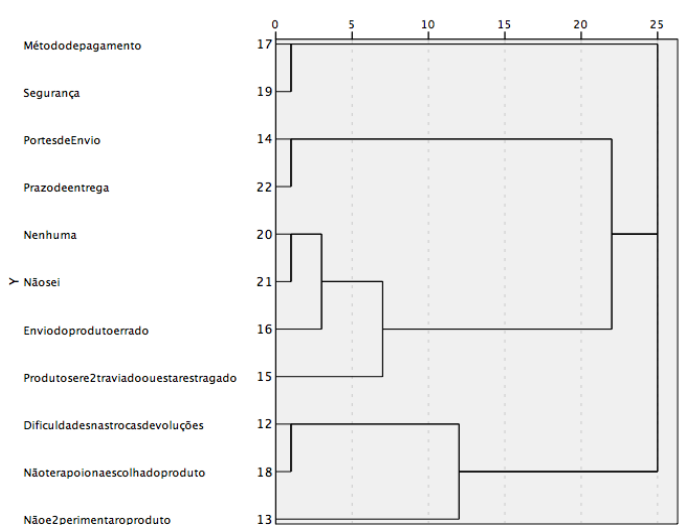
**Gráfico 46:** Dendograma das desvantagens usando o método de Ward para os inquiridos do sexo Masculino



**Gráfico 47:** Dendograma das desvantagens usando o método de Ward para os inquiridos não compradores



**Gráfico 48:** Dendograma das desvantagens usando o método de Ward para os inquiridos compradores



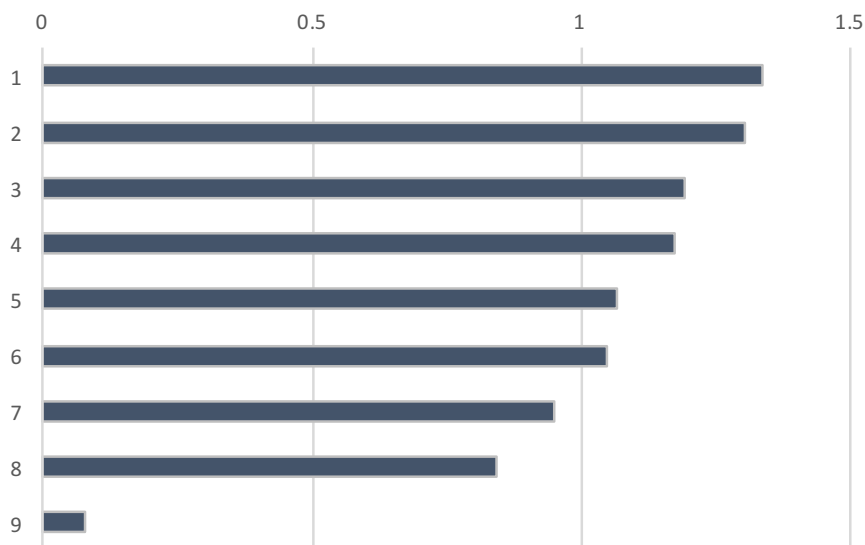
De seguida, e de modo a avaliar as diferenças entre as desvantagens identificadas entre compradores e não compradores e entre géneros relativamente às desvantagens da compra *Online* será feita uma comparação entre a Análise de Correspondências Múltiplas, a Análise Discriminante e a regressão logística, usando como variáveis iniciais “método de pagamento”, “segurança”, “envio do produto errado”, “produto ser extraviado ou estar estragado”, “não experimentar o produto”, “portes de envio”, “prazo de entrega”, “falta de apoio no pré e pós venda” e “não sei/nenhuma”.

### **Análise de correspondências múltiplas:**

Recorremos à ACM para realizar uma análise conjunta das desvantagens identificadas na compra online. Simultaneamente verificaremos se existem diferenças entre as desvantagens identificadas pelos compradores e pelos não compradores, bem como se existem diferenças entre géneros. Começou-se por se definir o número máximo de dimensões que permitem representar o conjunto de dados, i.e., 9 (anexo 7). Em função dos resultados obtidos, um critério passível de ser seguido para a escolha do número de dimensões seria a escolha das dimensões cujo valor próprio superior a 1 ou ainda optar-se por um número suficiente de dimensões que explicasse 80% da variância total. Seguindo estes métodos a opção passaria por escolher 6 dimensões, o que representaria uma redução fraca do número de variáveis iniciais.

Assim, e conjugando o decaimento dos valores próprios (Gráfico 49) e da análise das medidas de discriminação obtidas (Tabela 43) optou-se por 4 dimensões, uma vez que a partir da 5 dimensão, as variáveis relevantes para cada dimensão não adicionam nada de novo face às anteriores. Assim, o modelo escolhido representa 55,7% da variância total das variáveis.

**Gráfico 49:** Representação da variância das dimensões associadas às desvantagens da compra *online*



**Tabela 43:** Medidas de discriminação associadas à ACM das desvantagens da compra *online*

	Dimension				Mean
	1	2	3	4	
Não experimentar o produto	<b>.365</b>	.084	.055	<b>.333</b>	.209
Portes de Envio	<b>.296</b>	.135	.025	.087	.136
Produto ser extraviado ou estar estragado	.021	.026	<b>.368</b>	.178	.148
Envio do produto errado	.003	.062	.014	<b>.246</b>	.081
Método de pagamento	<b>.291</b>	.057	<b>.215</b>	.008	.143
Segurança	<b>.220</b>	.110	.170	<b>.253</b>	.188
Prazo de entrega	.059	<b>.200</b>	<b>.240</b>	.010	.127
Falta de apoio no pré e pós-venda	.073	<b>.598</b>	.106	.002	.195
Nenhuma/Não sei	.012	.034	.000	.057	.026
Costuma realizar compras Online?	.026	.002	.014	.009	.013
Sexo	.000	.002	.003	.003	.002
Active Total	1.339	1.306	1.195	1.175	1.254
% of Variance	14.882	14.516	13.272	13.058	13.932

Com base nos valores das medidas de discriminação e considerando as variáveis que mais contribuem dimensões, passaremos a denominar cada uma das dimensões anteriores do seguinte modo:

- Dimensão 1: **“Pagamentos e confiança”** (variáveis “Não experimentar o produto ao vivo”, “Portes de envio”, “Método de pagamento” e “Segurança”)
- Dimensão 2: **“Pós-venda”** (variáveis: “Falta de apoio no pré e pós venda” e “Prazo de Entrega”)
- Dimensão 3: **“Entrega e método de pagamento”** (variáveis “Produto ser extraviado ou estar estragado”, “método de pagamento” e “prazo de entrega”)
- Dimensão 4: **“Segurança e preocupação sobre recepção do produto pretendido”** (variáveis: “Não experimentar o produto ao vivo”, “Envio do produto errado” e “Segurança”)

Tal como anteriormente, os valores dos coeficientes de discriminação das variáveis “Costuma realizar compras *online*?” e “Sexo” são bastante próximas de zero para todas as dimensões, o que indicia a não existência de diferenças entre os elementos destes grupos.

**Tabela 44:** Posicionamento das variáveis mais relevantes para a ACM das desvantagens de compra *online*

	Quantificações negativas	Quantificações positivas
<u>Pagamentos e segurança/confiança</u> (Dimensão 1)	Não experimentar o produto_Sim Portes de Envio_Sim Método de pagamento_Não Segurança_Não	Não experimentar o produto_Não Portes de Envio_Sim Método de pagamento_Sim Segurança_Sim
<u>Pós-venda</u> (Dimensão 2)	Falta de apoio no pré e pós-venda_Não Prazo de entrega_Sim	Falta de apoio no pré e pós-venda_Sim Prazo de entrega_Não
<u>Entrega e método de pagamento</u> (Dimensão 3)	Produto ser extraviado ou estar estragado_Sim Método de pagamento_Não Prazo de entrega_Não	Método de pagamento_Sim Método de pagamento Prazo de entrega_Sim
<u>Segurança e preocupação sobre recepção do produto pretendido</u> (Dimensão 4)	Não experimentar o produto_Sim Envio do produto errado_Sim Segurança_Sim	Não experimentar o produto_Não Envio do produto errado_Não Segurança_Não

Assim, podemos interpretar a **Pagamentos e segurança/confiança** cujos indivíduos que se posicionam de modo oposto relativamente à segurança e ao método de pagamento relativamente ao Não experimentar o produto e aos portes de envio.

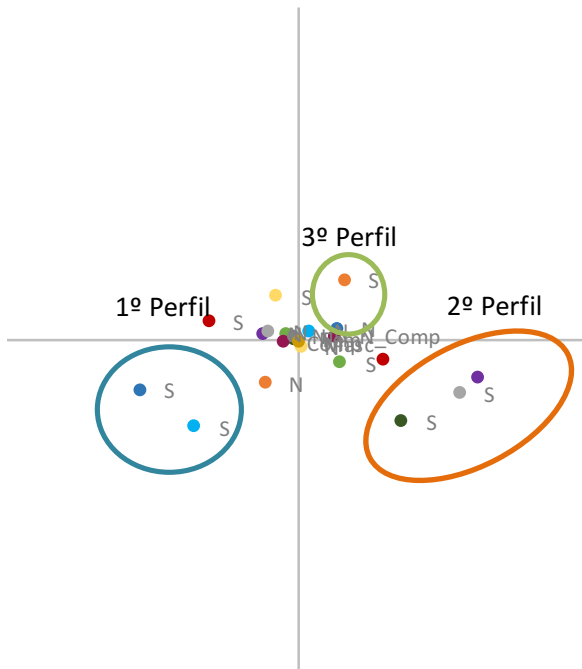
Já a dimensão de **Pós-venda** caracteriza-se pelos indivíduos que são opostos na falta de apoio no pré e no pós-venda relativamente ao prazo de entrega.

A dimensão de **Entrega e método de pagamento** é a dimensão na qual os indivíduos concordam relativamente ao método de pagamento e ao prazo de entrega, porém discordam face ao ao produto ser extraviado ou estar estragado.

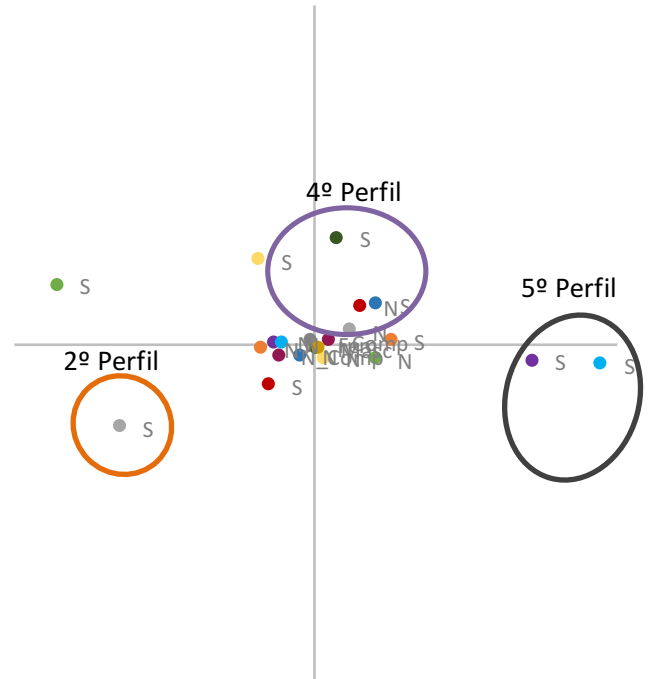
Finalmente, a dimensão **Segurança e preocupação sobre recepção do produto pretendido**, representa os indivíduos que concordam entre si com não experimentar o produto, o envio do produto errado e a segurança.

Com base no valor das coordenadas dos centroides, representou-se graficamente as dimensões mais relevantes quando comparadas duas a duas:

**Gráfico 50:** Coordenadas dos centroides da ACM para as desvantagens da compra *Online*:  
**Dimensão 1 vs Dimensão 2**



**Gráfico 51:** Coordenadas dos centroides da ACM para as desvantagens da compra *Online*:  
**Dimensão 3 vs Dimensão 4**



- |   |                                     |
|---|-------------------------------------|
| ● Não experimentar o produto                | ● Portes de Envio                   |
| ● Produto ser extraviado ou estar estragado | ● Envio do produto errado           |
| ● Método de pagamento                       | ● Segurança                         |
| ● Prazo de entrega                          | ● Falta de apoio no pré e pós-venda |
| ● Nenhuma/Não sei                           | ● Costuma realizar compras Online?  |
| ● Sexo                                      |                                     |

Da análise dos gráficos anteriores é possível identificar, tendo em conta as variáveis que mais discriminam as respectivas dimensões, 5 perfis distintos de indivíduos:

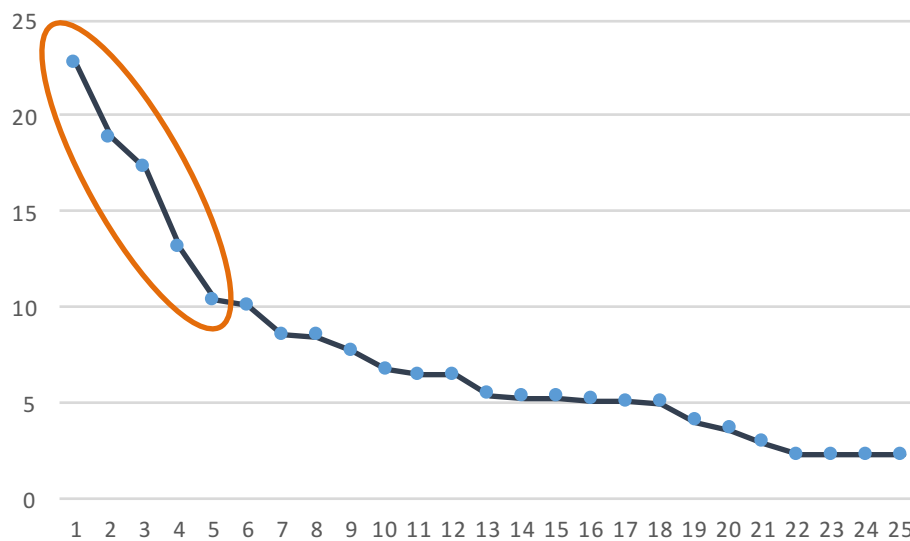
1. Indivíduos que consideram que as principais desvantagens associadas à compra *online* são o prazo de entrega, os portes de envio e o não experimentar o produto;
2. Indivíduos para os quais a segurança é o que os preocupa neste tipo de compra;
3. Indivíduos cuja principal preocupação é falta de apoio no pré e pós venda;
4. Indivíduos preocupados com o extravio do produto ou o produto estragado estar estragado;



5. Indivíduos para os quais as principais desvantagens são o método de pagamento e o prazo de entrega.

Destaca-se ainda que as coordenadas dos centroides para os dois géneros, se encontram bastante próximas entre si (e da origem), o que, mais uma vez indica, a não existência de diferenças entre os dois grupos no que diz respeito às desvantagens da compra *online*.

**Gráfico 52:** Coeficientes de fusão da análise de Clusters associada às desvantagens da compra *online*



A Análise de Clusters hierárquica sugere o agrupamento dos indivíduos em 5 grupos (Gráfico 53). Recorrendo à classificação não hierárquica (K-means) e cruzando com os perfis previamente identificados com a ACM, podemos dizer que as principais características dos grupos identificados são (Tabela 45):

- Grupo 1 (16% inquiridos) caracteriza-se pela preocupação com poder experimentar o produto ao vivo, os portes de envio, o prazo de entrega e a falta de apoio no pré e pós-venda;
- Grupo 2 (36% inquiridos) abrange o segmento dos indivíduos essencialmente preocupados com a segurança e que desvalorizam a questão do apoio no pré e pós venda;
- Grupo 3 (25% dos inquiridos) identifica o segmento dos que consideram que a falta de apoio quer no pré como no pós venda;
- Grupo 4 (15 % inquiridos) é caracterizado pelos indivíduos que tem receios relativamente à recepção do produto;
- Grupo 5 (8% dos inquiridos) é o de menor dimensão e refere-se a um segmento que considera que as principais desvantagens são o método de pagamento e o prazo de entrega, desvalorizando não experimentar o produto e a falta de apoio na venda.

**Tabela 45:** Caracterização dos 5 grupos identificados, relativamente aos às desvantagens da compra *online* identificadas pelos inquiridos.

	Categoria	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Total
Não experimentar o produto	Não	3,0%	66,7%	45,1%	<b>100,0%</b>	<b>100,0%</b>	50,7%
	Sim	<b>97,0%</b>	33,3%	54,9%	0,0%	0,0%	49,3%
Portes de Envio	Não	65,3%	98,3%	89,6%	76,9%	53,3%	80,0%
	Sim	<b>34,7%</b>	1,7%	10,4%	23,1%	<b>46,7%</b>	20,0%
Produto ser extraviado ou estar estragado	Não	78,2%	85,0%	100,0%	32,3%	100,0%	81,3%
	Sim	21,8%	15,0%	0,0%	<b>67,7%</b>	0,0%	18,8%
Envio do produto errado	Não	82,2%	100,0%	90,3%	72,3%	100,0%	87,5%
	Sim	17,8%	0,0%	9,7%	<b>27,7%</b>	0,0%	12,5%
Método de pagamento	Não	<b>100,0%</b>	83,3%	77,1%	87,7%	60,0%	84,3%
	Sim	0,0%	16,7%	22,9%	12,3%	<b>40,0%</b>	15,8%
Segurança	Não	100,0%	0,0%	100,0%	100,0%	100,0%	85,0%
	Sim	0,0%	<b>100,0%</b>	0,0%	0,0%	0,0%	15,0%
Prazo de entrega	Não	85,1%	98,3%	100,0%	96,9%	20,0%	89,5%
	Sim	<b>14,9%</b>	1,7%	0,0%	3,1%	<b>80,0%</b>	10,5%
Falta de apoio no pré e pós-venda	Não	<b>100,0%</b>	<b>83,3%</b>	9,0%	<b>70,8%</b>	<b>80,0%</b>	58,5%
	Sim	0,0%	16,7%	<b>91,0%</b>	29,2%	20,0%	41,5%
Sexo	Feminino	100,0%	100,0%	100,0%	86,2%	100,0%	97,8%
	Masculino	0,0%	0,0%	0,0%	13,8%	0,0%	2,3%

Total	16%	36%	25%	15%	8%	100%
-------	-----	-----	-----	-----	----	------

Nota: A negrito identificaram-se as principais características dos grupo.

**Análise discriminante:**

Utilizando um procedimento análogo ao realizado anteriormente, começou-se por introduzir todas as variáveis para o modelo, introduzindo todas as desvantagens (usando facto de ser comprador ou não comprador *Online* como variável de agrupamento) e fazendo uma seleção manual das variáveis utilizando o critério de seleção de variáveis com base nos valores da matriz de estrutura ( $>0,200$ ) – anexo 8), tendo obtido:

**Tabela 46:** Modelo de Análise Discriminante para as desvantagens na compra *Online*, usando como Variável de grupo “Costuma realizar compras online”

	Wilks' Lambda	F	df1	df2	Sig.
Portes de Envio	.961	16.275	1	398	.000
Método de pagamento	.993	2.867	1	398	.091
Segurança	.990	3.963	1	398	.047
Prazo de entrega	.975	10.142	1	398	.002
Nenhuma/Não sei	.997	1.278	1	398	.259

Os valores da estatística Lambda de Wilks (próximos de 1 para quase todas as variáveis), parecem sugerir a não existência de diferenças entre os dois grupos (compradores e não compradores). De modo a verificar esta indicação estudou-se ainda as funções de Fisher associadas a este modelo (Tabela 47):

**Tabela 47:** Coeficientes da função de Fisher associada à Análise Discriminante para as desvantagens da compra *Online*, usando como variável de grupo “Costuma realizar compras online”

	Costuma realizar compras Online?	
	Não	Sim
Portes de Envio	12.227	13.276
Método de pagamento	12.217	11.810
Segurança	14.806	14.472
Prazo de entrega	14.508	15.650
Nenhuma/Não sei	55.054	54.362
(Constant)	-59.527	-60.402

Deste modo as funções de Fisher resultantes da análise discriminante são da forma:

$$D_{NComp}(x) = 12,227. portes\_de\_envio + 12,217. metodo\_de\_pagamento + 14,806. seguranca + 11,508. prazo\_de\_entrega + 55,054 nenhuma\_naosei - 59,527$$

$$D_{Comp}(x) = 13,276. portes\_de\_envio + 11,810. metodo\_de\_pagamento + 14,472. seguranca + 15,650. prazo\_de\_entrega + 54,362 nenhuma\_naosei - 60,402$$

Assim, e considerando as pequenas diferenças entre os coeficientes de cada uma das funções de Fisher, temos que não existem diferenças relevantes entre compradores e não compradores relativamente às desvantagens identificadas no *e-commerce*.

Através do modelo apresentado, podemos ainda calcular a percentagem de indivíduos bem classificados em cada um dos grupos (compradores ou não compradores), sendo que, como é possível verificar, 90,1% dos não compradores são bem classificados, porém apenas 35,3% dos compradores são bem classificados através deste modelo (Tabela 48). Verifica-se assim que o modelo apresenta deficiências na classificação dos compradores.

**Tabela 48:** Probabilidade de um novo elemento pertencer ao grupo dos compradores, através do modelo obtido para as desvantagens da compra online

		Costuma realizar compras Online?	Predicted Group Membership		Total
			Não	Sim	
Original	%	Não	90.1	9.9	100.0
		Sim	64.7	35.3	100.0

a. 50.5% das observações iniciais foram corretamente classificadas.

Utilizando um raciocínio análogo, utilizando a variável sexo como *variável de agrupamento*, e usando o mesmo critério para escolha das variáveis, variáveis, obtiveram-se os resultados apresentados na Tabela 49.

**Tabela 49:** Modelo de Análise Discriminante para as desvantagens na compra Online, usando como variável de grupo "Sexo"

	Wilks' Lambda	F	df1	df2	Sig.
Não experimentar o produto	.995	1.865	1	398	.173
Produto ser extraviado ou estar estragado	.978	9.066	1	398	.003

Método de pagamento	.994	2.307	1	398	.130
Prazo de entrega	.985	5.974	1	398	.015
Nenhuma/Não sei	.996	1.477	1	398	.225

Tal como no exemplo anterior, os valores da estatística de Lambda de Wilks (próximos de 1 para todas as variáveis), parecem sugerir a não existência de diferenças entre os dois géneros. De modo a verificar esta indicação estudou-se ainda as funções de Fisher associadas a este modelo

**Tabela 50:** Coeficientes da função de Fisher associada à Análise Discriminante para as desvantagens da compra *Online*, usando como variável de grupo "Sexo"

	Sexo	
	Feminino	Masculino
Não experimentar o produto	15.813	15.577
Produto ser extraviado ou estar estragado	17.234	17.996
Método de pagamento	18.640	18.219
Prazo de entrega	18.360	19.257
Nenhuma/Não sei	62.826	62.012
(Constant)	-75.750	-75.991

Assim, as funções de Fisher resultantes da análise discriminante são da forma:

$$\begin{aligned}
 D_{Fem}(x) = & 15,813.nao\_experimentar\_o\_produto \\
 & + 17,234.produto\_ser\_extraviado\_ou\_estar\_estragado \\
 & + 18,640.metodo\_de\_pagamento + 18,360.prazo\_de\_entrega \\
 & + 62,826nenhuma\_nãosei - 75,750
 \end{aligned}$$

$$\begin{aligned}
 D_{Masc}(x) = & 15,577.nao\_experimentar\_o\_produto \\
 & + 17,996.produto\_ser\_extraviado\_ou\_estar\_estragado \\
 & + 18,219.metodo\_de\_pagamento + 19,257.prazo\_de\_entrega \\
 & + 62,012nenhuma\_nãosei - 75,991
 \end{aligned}$$

Assim e considerando as pequenas diferenças entre os coeficientes de cada uma das funções de Fisher temos que não existem indícios de diferenças relevantes entre as desvantagens identificadas pelos dois géneros, confirmando as orientações do teste de Wilks.

Calculou-se, ainda, através da classificação efetuada probabilidade de um novo elemento da amostra pertencer a um dos dois grupos (homens ou mulheres), sendo que, como é possível verificar, 78,9% dos elementos do sexo feminino são bem classificados, porém apenas 39,3% dos homens são bem classificados através deste modelo. Deste modo, constata-se que este modelo apresenta deficiências na classificação correcta dos homens.

**Tabela 51:** Percentagem de homens e mulheres classificados corretamente, através do modelo obtido para as desvantagens associadas às compras *online*

		Sexo	Predicted Group Membership		Total
			Feminino	Masculino	
Original	%	Feminino	78.9	21.1	100.0
		Masculino	60.7	39.3	100.0

a. 62.3% das observações iniciais foram corretamente classificadas.

### Regressão Logística:

Para construir um modelo de regressão logística para a variável “Sexo” considerando as desvantagens na compra *online* como variáveis independentes, introduziram-se todas as desvantagens em simultâneo e foi feita uma seleção manual das variáveis utilizando o índice  $\left| \frac{S.E.}{\beta} \right|$ , sendo que quanto menores forem os valores deste rácio, maior será a contribuição da variável para o modelo. Utilizando este raciocínio, chegou-se ao seguinte modelo apresentado na Tabela 52, do qual é possível verificar que os portes de envio e o prazo de entrega são desvantagens importantes relativamente a comprar *online*.

**Tabela 52:** Modelo de Regressão Logística para os compradores *online* em função das desvantagens na compra *online*

	$\beta$	S.E.	$\left  \frac{S.E.}{\beta} \right $	$\text{Exp}(\beta)$
Portes de Envio	1.701	.405	0,24	5.480
Envio do produto errado	.524	.358	0,68	1.689
Prazo de entrega	2.029	.624	0,31	7.606
Falta de apoio no pre e pos-venda	.620	.246	0,40	1.859

Constant	-4.629	1.104	0,24	.010
----------	--------	-------	------	------

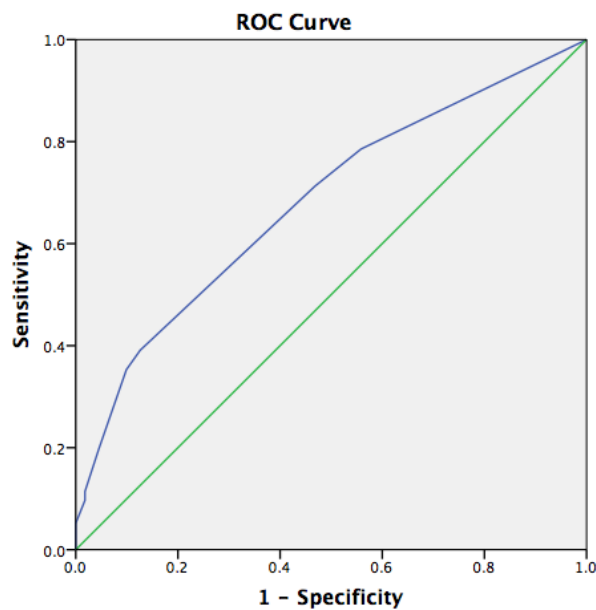
Para um ponto de corte de 0,70 obtemos um modelo com sensibilidade igual a 71,3% e uma especificidade igual a 53,2%, isto é são identificados corretamente 71,3% dos compradores e 53,2% dos não compradores, concluindo que o modelo ajusta corretamente 66,3% dos inquiridos o que representa uma qualidade razoavel de ajustamento. À semelhança do verificado com a Análise discriminante, também este odelo apresenta deficiências na classificação correcta dos compradores. O modelo encontrado tem uma fraca capacidade discriminativa (AUC=0,681).

**Tabela 53:** Valores previstos com o Modelo de Regressão Logística obtido para os compradores em função das desvantagens, usando o ponto de corte 0,70

Observed		Predicted		
		Costuma realizar compras Online?		Percentage Correct
		Não	Sim	
Costuma realizar compras Online?	Não	59	52	53.2
	Sim	83	206	71.3
Overall Percentage				66.3

a. Ponto de corte = .700

**Gráfico 53:** Curva ROC associada ao Modelo de Regressão Logística obtido para as desvantagens na compra Online, usando como variável suplementar “Costuma realizar compras online”



Diagonal segments are produced by ties.

Através dos valores  $\hat{\beta}$ , podemos construir a função logística associada ao modelo obtido:

$$\begin{aligned} \text{logit}(\hat{\pi}(x)) &= \beta_0 + \beta_1.\text{portes\_de\_envio} + \beta_2.\text{envio\_do\_produto\_errado} + \beta_3.\text{prazo\_de\_entrega} \\ &\quad + \beta_4.\text{falta\_de\_apoio\_no\_pre\_e\_posvenda} \\ &= 1.701.\text{portes\_de\_envio} + 0,524.\text{envio\_do\_produto\_errado} + 2,029.\text{prazo\_de\_entrega} \\ &\quad + 0,620.\text{falta\_de\_apoio\_no\_pre\_e\_posvenda} - 4,629 \end{aligned}$$

Repetindo o mesmo raciocínio, agora para modelar o sexo em função das desvantagens identificadas na compra *online*, o prazo de entrega é um factores potenciadores para ser do sexo masculino (Tabela 54).

**Tabela 54:** Modelo de Regressão Logística obtido para as desvantagens na compra *Online*, usando como variável suplementar “Costuma realizar compras *online*”

	B	S.E.	$\left  \frac{S.E.}{\beta} \right $	Exp(B)
Produto ser extraviado ou estar estragado	.989	0,237	0,24	2.688
Prazo de entrega	1.149	0,781	0,68	3.157
Falta de apoio no pre e pos-venda	.395	0,122	0,31	1.484
Segurança	.435	0,174	0,40	1.544
Constant	-3.838	0,921	0,24	.022

Assim, ao utilizarmos um ponto de corte de 0,40 obtemos um modelo com sensibilidade igual a 42,3% e uma especificidade igual a 76,7%, isto é, são identificados corretamente 42,3% dos indivíduos do sexo masculino e 76,7% dos indivíduos do sexo feminino, concluindo que o modelo ajusta corretamente 62,3% dos inquiridos o que representa uma boa qualidade de ajustamento. Porém, o modelo encontrado tem ainda uma fraca capacidade discriminativa (AUC = 0,615).

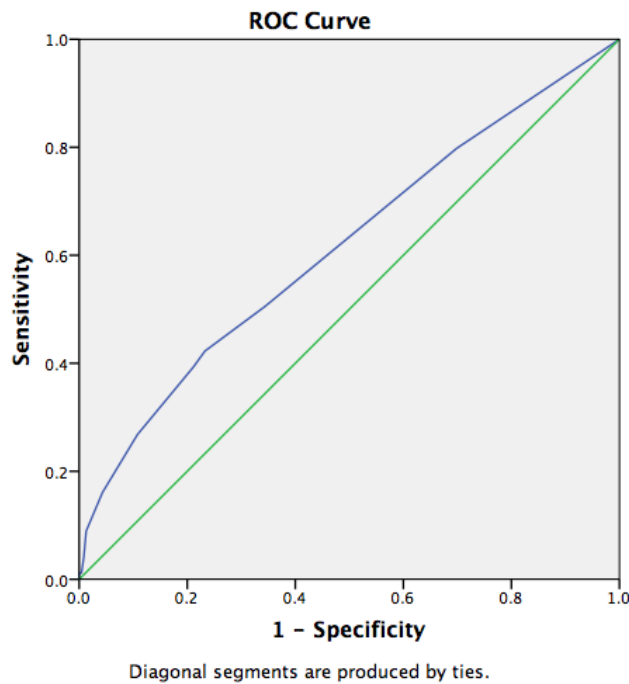


**Tabela 55:** Valores previstos com o Modelo de Regressão Logística obtido para os sexos em função das desvantagens, usando o ponto de corte 0,40

Observed		Predicted		
		Sexo		Percentage Correct
		Feminino	Masculino	
Sexo	Feminino	178	54	76,7
	Masculino	97	71	42,3
Overall Percentage				62,3

a. Ponto de corte = 0,400

**Gráfico 54:** Curva ROC associado ao Modelo de Regressão Logística obtido para as desvantagens na compra Online, usando como variável suplementar “Sexo”



Através dos valores  $\hat{\beta}$ , podemos construir a função logística associada ao modelo obtido:

$$\begin{aligned} \text{logit}(\hat{\pi}(x)) &= \beta_0 + \beta_1.\text{produto_ser_extraviado_ou_estar_estragado} + \beta_2.\text{prazo_de_entrga} \\ &\quad + \beta_3.\text{falta_de_apoio_no_pre_e_posvenda} + \beta_4.\text{seguranca} \\ &= 0,989.\text{produto_ser_extraviado_ou_estar_estragado} + 1,149.\text{prazo_de_entrga} \\ &\quad + 0,395.\text{falta_de_apoio_no_pre_e_posvenda} + 0,435.\text{seguranca} - 3,838 \end{aligned}$$

Por via a comparar as variáveis e conclusões retiradas através da Análise de Correspondências múltipla, da função de Fisher resultante da Análise Discriminante e da Regressão Logística quando sujeitas ao mesmo problema, concluiu-se que quer na Análise de Correspondências Múltiplas (ACM) quer através da Função de Fisher resultante da Análise Discriminante (AD) conclui-se que não existem diferenças entre compradores e não compradores e entre géneros no que diz respeito às desvantagens associadas à compra *online* identificadas pelos indivíduos na amostra em estudo. No que diz respeito às variáveis identificadas como relevantes na Análise Discriminante (AD) e na Regressão Logística (RL), registou-se que:

- As variáveis escolhidas, para AD para os grupos construídos por homens/mulheres e compradores/não compradores, foram as mesmas;
- Existiram diferenças entre as variáveis selecionadas para o modelo de RL entre os dois grupos;
- Os modelos de AD e RL identificaram variáveis distintas no estudo para os grupos de compradores/não compradores e homens/mulheres;

**Tabela 56:** Tabela resumo dos métodos utilizados na análise das desvantagens em comprar *Online*

	Compradores vs. Não Compradores		Homens vs. Mulheres	
	Existem indícios de diferenças entre os grupos?	Variáveis diferenciadoras	Existem indícios de diferenças entre os grupos?	Variáveis diferenciadoras
Análise de Correspondências Múltiplas	Não		Não	
Análise Discriminante -> Função de Fisher	Não		Não	
Regressão Logística	Sim	<ul style="list-style-type: none"> <li>• Portes de envio</li> <li>• Envio do produto errado</li> <li>• Prazo de entrega</li> <li>• Falta de apoio no pré e pós-venda</li> </ul>	Sim	<ul style="list-style-type: none"> <li>• Falta de apoio no pré e pós-venda</li> <li>• Segurança</li> </ul>

Quando questionados sobre quais os motivos que poderiam levar a comprar/comprar mais online, não houve diferenças evidentes entre géneros, porém detetaram-se algumas discrepâncias entre as opiniões de compradores e não compradores (

### 4.6.3) Motivações e preferências na compra *Online*

Tabela 57). Os compradores comprariam mais online caso exista opção de entrega rápida (32,8% face a não compradores), pressão de familiares/amigos (45,3% contra 39,8%) ou portes gratuitos (15,3% versus a 12,7% dos não compradores). Já os não compradores comprariam online poderiam tornar-se compradores caso existam descontos exclusivos online (8,5% face a 1,1% dos compradores), caso o processo de compra seja simples e rápido (7,6% versus 1,8%) e ainda por outros motivos não especificados (3,4% face a 0,4% dos compradores).

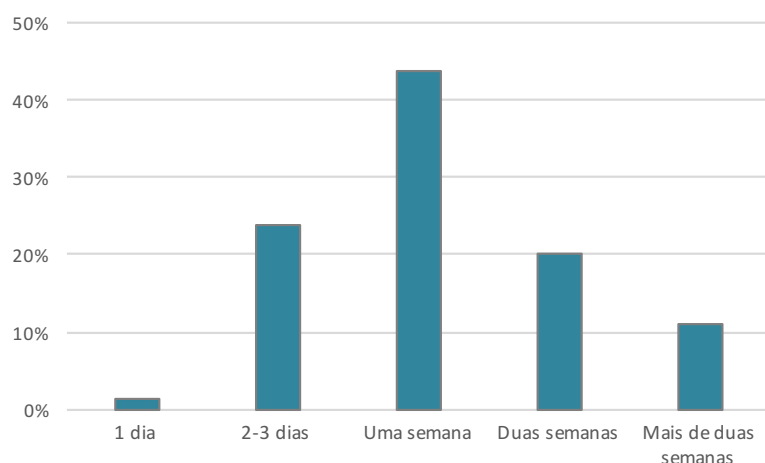
### 4.6.4) Motivações e preferências na compra *Online*

**Tabela 57:** Quais os motivos que o poderiam levar a comprar/comprar mais *online*?

	Género		Costuma realizar compras <i>online</i> ?	
	Masculino	Feminino	Compradores	Não compradores
Portes de Envio Gratuitos	17,2%	12,7%	15,3%	12,7%
Descontos Exclusivos Online	3,1%	3,5%	1,1%	8,5%
Opção de entrega rápida	30,7%	30,1%	32,8%	24,6%
Não sei	0,0%	0,4%	0,4%	0,0%
Processo de compra simples e rápido	4,9%	2,6%	1,8%	7,6%
Produto situado fora do alcance habitacional	0,6%	0,4%	0,7%	0,0%
Outros	1,8%	0,9%	0,4%	3,4%
Outros modos de pagamento	0,6%	0,0%	0,4%	0,0%
Possibilidade de troca/devolução sem custos	0,6%	3,5%	1,8%	3,4%
Pressão de amigos/familiares	40,5%	45,9%	45,3%	39,8%

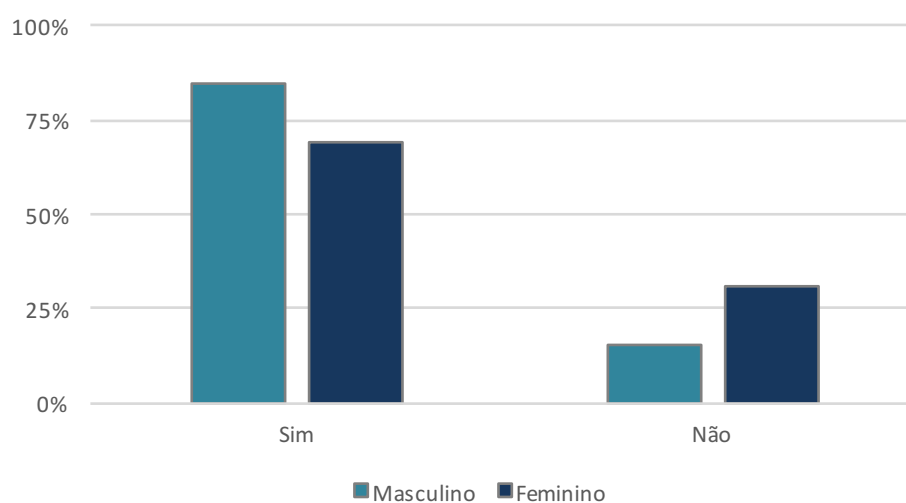
Outro dado a reter é que apenas 1,5% dos inquiridos exige receber o produto comprado *online* no dia seguinte, enquanto que 11,0% está disposto a esperar mais de duas semanas pela entrega da compra efetuada.

**Gráfico 55:** Tempo máximo para entrega de compra *Online*



Uma das principais vantagens das lojas *online* para os consumidores está relacionada com o utilizador poder escolher e comparar os produtos entre várias lojas/fornecedores em simultâneo. Neste sentido foi questionado aos inquiridos se estes costumam comparar produtos *online*, sendo que 75,8% dos inquiridos responderam afirmativamente, sendo esta percentagem superior nos homens (84,5% face a 69,4% das mulheres inquiridas) e, naturalmente, entre os compradores *online* (94,1%).

**Gráfico 56:** Inquiridos que comparam produtos *Online*



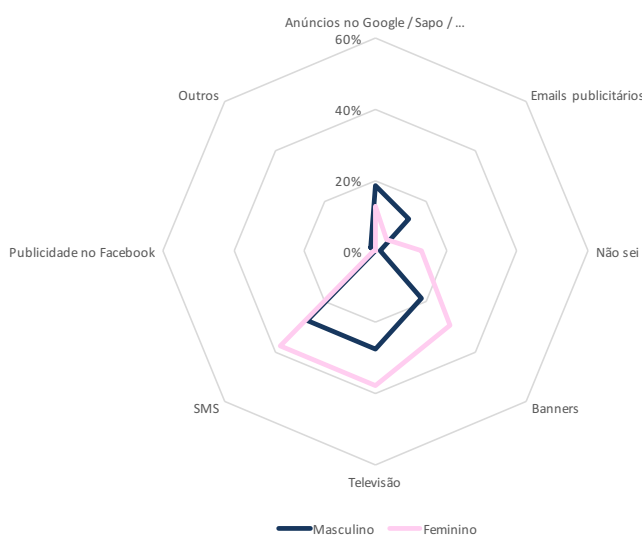
#### 4.6.4) Meios de comunicação *Online*

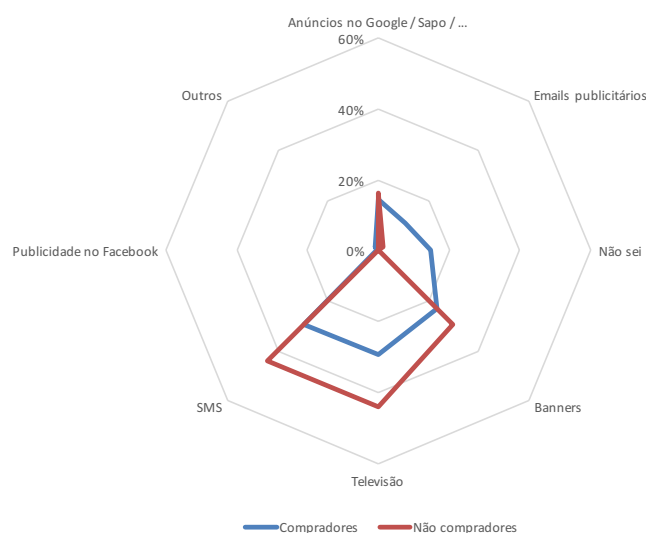
No que diz respeito aos meios de comunicação considerados mais eficazes para publicitar um produto *online*, o meio preferencial para 48,75% dos inquiridos é o email. De destacar ainda a publicidade no Facebook (33,50%) e ainda os anúncios de texto em motores de busca (25,50%). Uma percentagem ainda relevante dos inquiridos (15,25%) revelou não saber qual o meio que considera mais eficaz.

Os inquiridos do sexo feminino mostraram preferir mais os emails publicitários (51,29% face a 45,24% dos homens), e a publicidade no Facebook (37,93% versus 27,38%). Já os homens mostraram preferir todos os restantes meios de angariação, com destaque para os *banners* de publicidade (13,10% versus 4,74%). Note-se que foram também os homens os que mais denotaram desconhecimento sobre o meio de comunicação mais eficaz: 18,45% versus 12,93% das inquiridas.

Por via a avaliar se as diferenças encontradas relativamente os meios de publicidade que compradores e não compradores e homens e mulheres identificaram como aqueles que garantem maior sucesso junto dos utilizadores optou-se por comparar novamente os resultados da Análise de Correspondências Múltiplas, da função de Fisher resultante da Análise Discriminante e da Regressão Logística.

**Gráfico 57:** Meios de publicidade considerados mais eficazes por cada um dos sexos

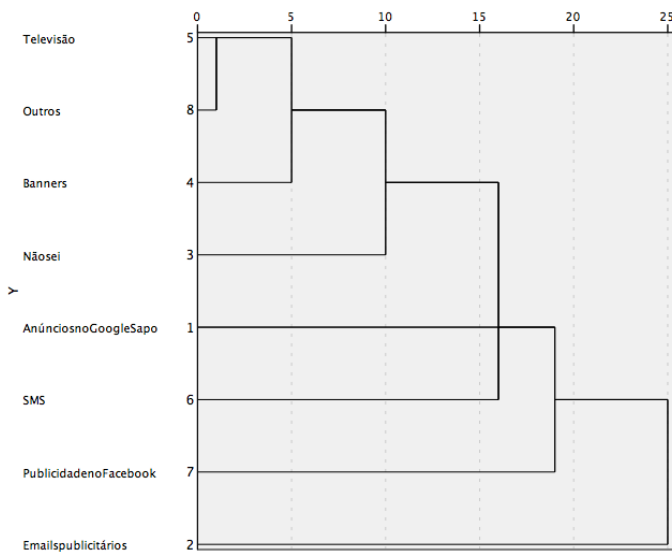


**Gráfico 58:** Meios de publicidade considerados mais eficazes por Compradores e Não Compradores *Online*

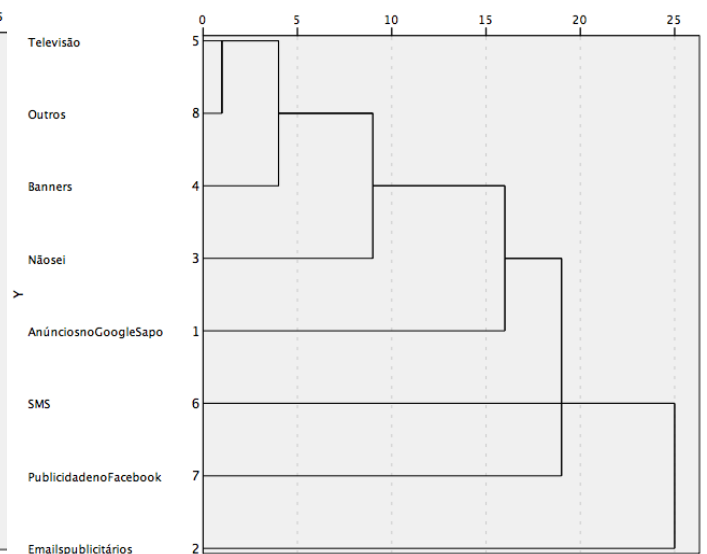
De modo a avaliar as diferenças entre os meios de publicidade mais eficazes identificados por compradores e não compradores e entre géneros, serão comparados três métodos: a Análise de Correspondências Múltiplas, a análise da função de Fisher associada à Análise Discriminante e a Regressão Logística. Porém, uma vez que existem variáveis com um número reduzido de observações, optou-se por se efetuar uma Análise de Clusters, com carácter exploratório, de modo a identificar eventuais associações entre as variáveis por via a agrupar variáveis com poucas observações.

Numa primeira fase traçou-se uma análise de *Clusters* para a globalidade dos dados, usando vários métodos de classificação hierárquica: centroides, maior distância e método de *Ward*, usando como medida de associação o Phi, tendo-se obtido os seguintes dendogramas e respetivo gráficos dos coeficientes de fusão.

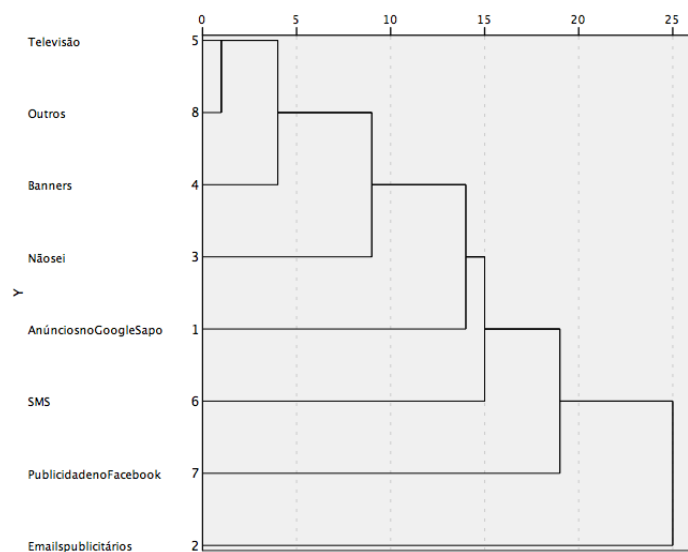
**Gráfico 59:** Dendograma dos meios de publicidade mais eficazes usando o método dos centroides



**Gráfico 60:** Dendograma publicidade mais eficazes usando o método da maior distância



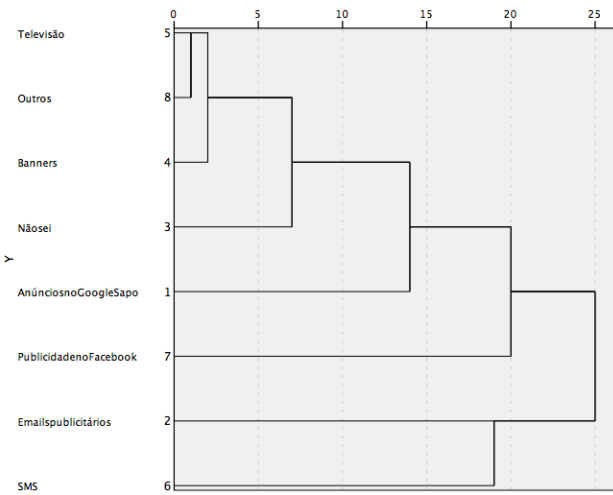
**Gráfico 61:** Dendograma publicidade mais eficazes usando o método de Ward



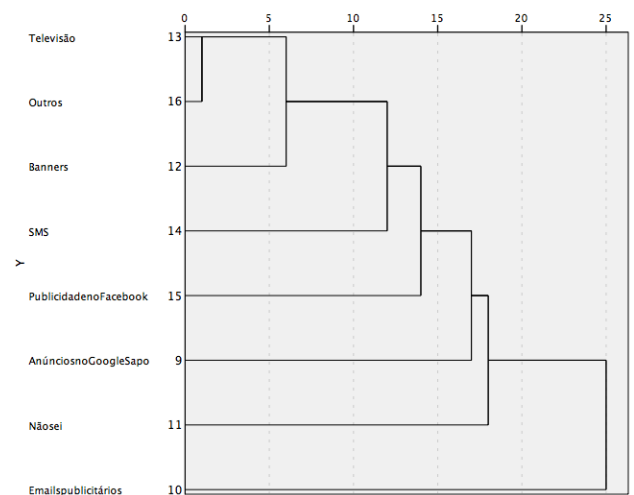
Através da análise dos dendogramas anteriores, verificou-se que todos os métodos possuem um bom poder discriminativo das variáveis. Os gráficos dos coeficientes de fusão não se mostraram eficazes no agrupamento das variáveis. Assim, optou-se por unificar as variáveis “Televisão” e “Outros” (que doravante será denominada por “Televisão/Outros”).

A escolha destas variáveis foi validada também pelos dendogramas efetuados através da análise de clusters usando as mesmas variáveis para ambos os sexos e para compradores e não compradores, de acordo com o método de Ward:

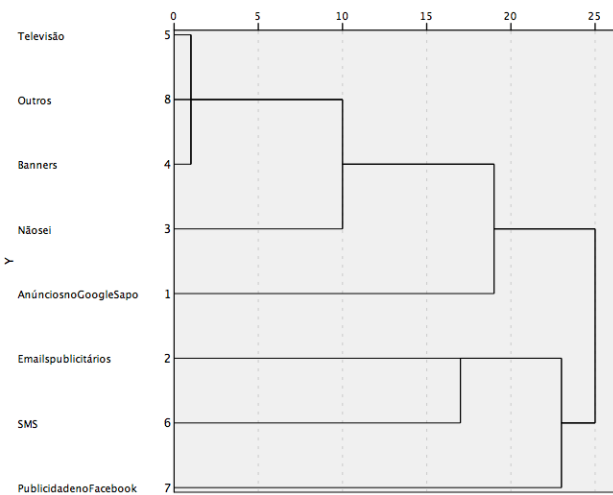
**Gráfico 62:** Dendograma das desvantagens usando o método de Ward para os inquiridos do sexo Feminino



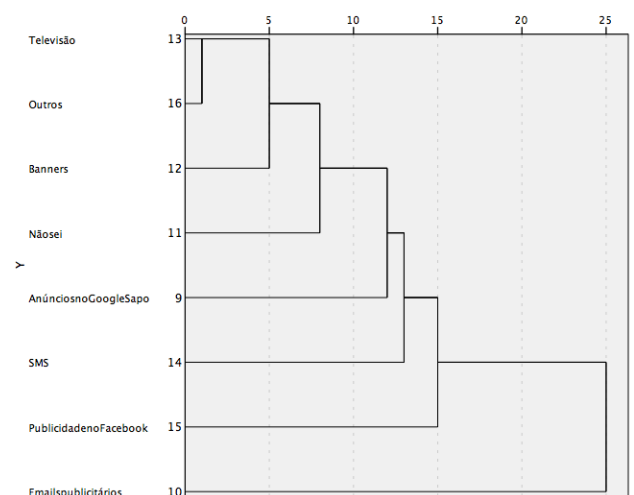
**Gráfico 63:** Dendograma das desvantagens usando o método de Ward para os inquiridos do sexo Masculino



**Gráfico 64:** Dendograma das desvantagens usando o método de Ward para os inquiridos não compradores



**Gráfico 65:** Dendograma das desvantagens usando o método de Ward para os inquiridos compradores



De seguida, e de modo a avaliar as diferenças entre as desvantagens identificadas entre compradores e não compradores e entre géneros relativamente às desvantagens da compra *Online* será feita uma



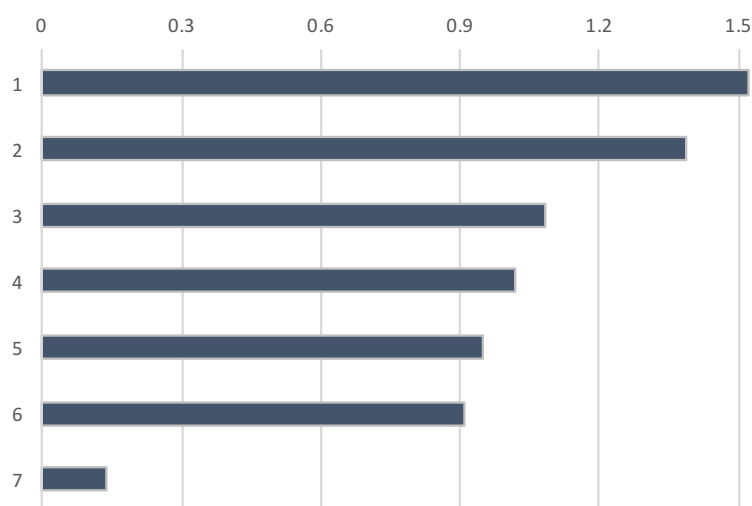
comparação entre a Análise de Correspondências Múltiplas, a Análise Discriminante e a regressão logística, usando como variáveis “Televisão/Outros”, “Banners”, “Não sei”, “Anúncios no Google / Sapo...”, “SMS”, “Publicidade no Facebook” e “Emails publicitários”.

### **Análise de correspondência múltiplas:**

Recorremos à ACM para realizar uma análise conjunta dos meios de publicidade mais eficazes identificados pelos inquiridos. Simultaneamente verificaremos se existem diferenças entre as meios de publicidade mais eficazes identificados por compradores e pelos não compradores, bem como se existem diferenças entre géneros. Começou-se por se definir o número máximo de dimensões que permitem representar o conjunto de dados, i.e., 7 (anexo 9). Em função dos resultados obtidos, um critério passível de ser seguido para a escolha do número de dimensões seria a escolha das dimensões cujo valor próprio superior a 1 ou ainda optar-se por um número suficiente de dimensões que explicasse 80% da variância total. Seguindo estes métodos a opção passaria por escolher 5 dimensões.

Assim, e conjugando o o decaimento dos valores próprios (Gráfico 66) e da análise das medidas de discriminação obtidas(Tabela 58) optou-se por 4 dimensões, uma vez que a partir da 5 dimensão, as variáveis relevantes para cada dimensão não adicionam nada de novo face às anteriores. Assim, o as dimensões retidas representam 71,6% da variância total.

**Gráfico 66:** : Representação da variância das dimensões associadas aos meios de publicidade mais eficazes



**Tabela 58:** Medidas de discriminação associadas à ACM dos meios de publicidade mais eficazes

	Dimension				Mean
	1	2	3	4	
Anúncios no Google / Sapo / ...	.082	<b>.430</b>	.060	.002	.143
Emails publicitários	<b>.587</b>	.008	.023	.002	.155
Não sei	<b>.536</b>	<b>.356</b>	.007	.039	.235
Banners	.011	.088	<b>.536</b>	.000	.159
SMS	.290	.128	.000	.001	.105
Publicidade no Facebook	.003	<b>.366</b>	<b>.458</b>	.000	.207
Televisão/Outros	.013	.009	.000	<b>.974</b>	.249
Costuma realizar compras Online? <sup>a</sup>	.004	.003	.036	.004	.012
Sexo <sup>a</sup>	.020	.000	.026	.003	.012
Active Total	1.523	1.385	1.084	1.018	1.252
% of Variance	21.750	19.785	15.487	14.543	17.891

a. Supplementary variable.

Com base nos valores das medidas de discriminação (Tabela 58) e considerando as variáveis que mais contribuem para a formação das dimensões, passaremos a denominar cada uma das dimensões anteriores do seguinte modo:

- Dimensão 1: **“Emails publicitários e não sei”** (variáveis “Emails publicitários” e “Não sei)
- Dimensão 2: **“Pesquisa e redes sociais”** (variáveis: “Anúncios no Google / Sapo /...”, “Não sei” e “Publicidade no Facebook”)
- Dimensão 3: **“Meios de publicidade de indução de procura”** (variáveis “Banners” e “Publicidade no Facebook” )
- Dimensão 4: **“Televisão/Outros”** (variável: “Televisão/Outros”)

Note-se ainda que os valores das medidas de discriminação das variáveis “Costuma realizar compras online?” e “Sexo” são bastante próximas de zero para todas as variáveis, o que indicia a não existência de diferenças entre os elementos destes grupos.

**Tabela 59:** Posicionamento das variáveis mais relevantes para a ACM dos meios de publicidade mais eficazes

	Quantificações negativas	Quantificações positivas
<u>Emails publicitários e não sabe</u> (Dimensão 1)	Emails publicitários_Não Não sei_Sim	Emails publicitários_Sim Não sei_Não
<u>Pesquisa e redes sociais</u> (Dimensão 2)	Anúncios no Google / Sapo / ..._Sim Não sei_Não Publicidade no Facebook_Sim	Anúncios no Google / Sapo / ..._Não Não sei_Sim Publicidade no Facebook_Não
<u>Meios de publicidade de indução de procura</u> (Dimensão 3)	Banners_Sim Publicidade no Facebook_Não	Banners_Não Publicidade no Facebook_Sim
<u>Televisão/Outros</u> (Dimensão 4)	Televisão/Outros_Não	Televisão/Outros_Sim

Com base na informação da Tabela 59 podemos interpretar a dimensão **Emails publicitários e não sabe** como a que caracteriza os indivíduos que se posicionam de modo oposto relativamente a Emails de publicidade e a não saber quais os meios de publicidade *online* mais eficazes.

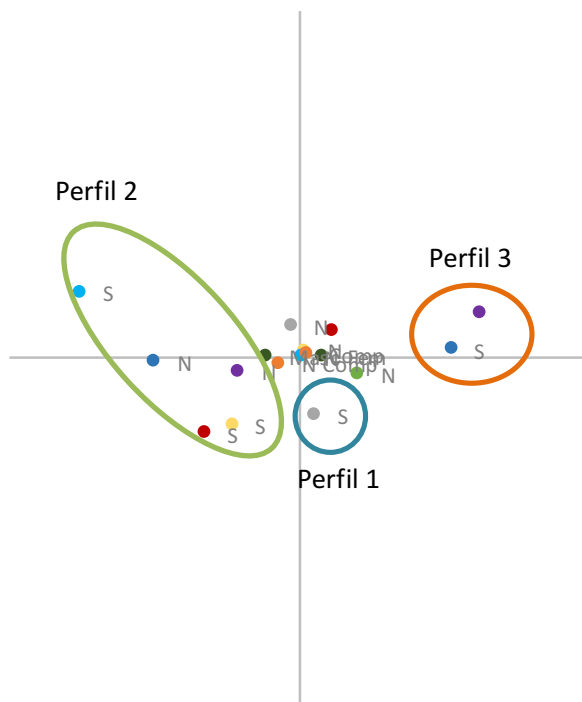
Já a dimensão de **Pesquisa e redes sociais** caracteriza-se pelos indivíduos que concordam relativamente a Anúncios no Google / Sapo / ...\_ e a Publicidade no Facebook, porém discordam em não saber quais os meios mais eficazes.

A dimensão de **Meios de publicidade de indução de procura** é a dimensão que opõe os indivíduos que consideram os *Banners* como meios eficazes de publicidade relativamente aos que consideram a Publicidade no Facebook.

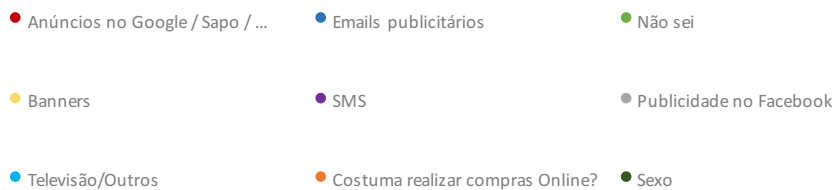
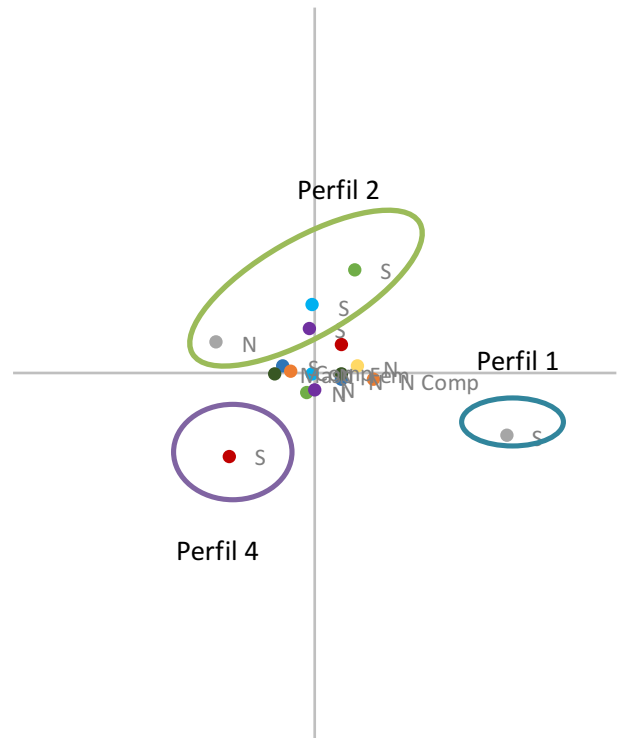
Finalmente, a dimensão **Televisão/Outros**, representa os indivíduos que consideram que os meios publicitários mais eficazes são a televisão e outros.

Com base no valor das coordenadas dos centroides, representou-se graficamente as dimensões mais relevantes quando comparadas duas a duas:

**Gráfico 67:** Coordenadas dos centroides da ACM para os meios de publicidade mais eficazes: **Dimensão 1 vs Dimensão 2**



**Gráfico 68:** Coordenadas dos centroides da ACM para os meios de publicidade mais eficazes: **Dimensão 3 vs Dimensão 4**



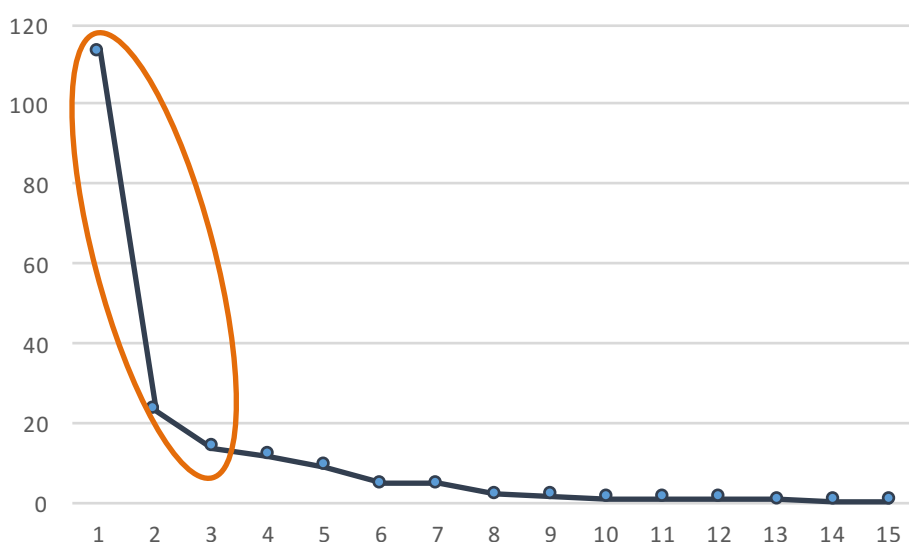
Da análise dos gráficos anteriores é possível identificar, tendo em conta as variáveis que mais discriminam as respectivas dimensões, 4 perfis distintos de compradores:

1. Indivíduos que preferem meios mais generalistas como televisão/outras, anúncios no Google/Sapo/... e Banners, por oposição às SMS e aos emails publicitários;
2. Indivíduos que o Facebook é o meio publicitário mais eficaz;
3. Indivíduos que consideram que meios mais reactivos como os banners e as SMS são os meios mais eficazes;
4. Indivíduos que preferem os anúncios no Google/Sapo/... a publicidade no Facebook.

Destaca-se ainda que as coordenadas dos centroides para os dois géneros, se encontram bastante próximas entre si (e da origem), o que, mais uma vez indica, a não existência de diferenças entre os grupos no que diz respeito aos meios publicitários mais eficazes.

A Análise de Clusters hierárquica aplicada às coordenadas dos indivíduos nas dimensões retidas sugere o agrupamento dos indivíduos em 3 grupos (Gráfico 69).

**Gráfico 69:** Coeficientes de fusão da análise de Clusters associada aos meios de publicidade mais eficazes – método do vizinho mais afastado



O gráfico anterior confirma a manutenção do critério anterior, isto é a divisão dos dados em 3 grupos.

Assim, e por via a enriquecer os resultados obtidos para compradores vs. Não compradores e por género recorreu-se base nos scores de cada observação em cada dimensão, obtidos na análise anterior utilizou-se o método k-means de forma a agrupar cada observação em 3 grupos, sendo que foi obtido que o grupo 1 seria composta por 193 indivíduos, o grupo 2 por 64 indivíduos e o grupo 3 por 143.

A partir deste clustering desenhou-se uma tabela de contigência por via a caracterizar o perfil dos inquiridos nos 3 grupos identificados, tendo-se concluído que o grupo 1 é caracterizado pelos indivíduos que têm uma opinião formada sobre os meios de publicidade mais eficazes, mas que porém não valorizam *banners* nem televisão/outros.

Os indivíduos do grupo 2 alegam desconhecimento dos meios de publicidade mais eficazes, não considerando nenhum dos apresentados como eficazes.

Já o terceiro grupo, também tem opinião formada sobre os meios publicitários mais relevantes, considerando que os emails de publicidade são os meios mais eficazes, contrastando com os anúncios no Google/sapo/..., os banners, as SMS e televisão/outros.

A partir da Análise de Clusters não hierárquica *K-means* é possível caracterizar os indivíduos de cada um dos grupos da seguinte forma (Tabela 60):

- Grupo 1 (48 % dos inquiridos): caracteriza-se pelos indivíduos que consideram os *banners* e a publicidade no facebook como os meios de publicidade mais eficazes;
- Grupo 2 (16 % dos inquiridos): é composto pelos indivíduos que preferem a televisão/outros face a todos os outros meios publicitários
- Grupo 3 (36 % dos inquiridos): representa os indivíduos que enaltecem os emails publicitários.

**Tabela 60:** Caracterização dos 3 grupos identificados, relativamente aos meios de publicidade mais eficazes, identificados pelos inquiridos.

	Categoria	Grupo 1	Grupo 2	Grupo 3	Total
Anúncios no Google / Sapo / ...	Não	50,0%	<b>96,8%</b>	85,1%	74,5%
	Sim	50,0%	3,2%	14,9%	25,5%
Emails publicitários	Não	<b>100,0%</b>	<b>100,0%</b>	0,0%	51,2%
	Sim	0,0%	0,0%	<b>100,0%</b>	48,8%
Não sei	Não	<b>100,0%</b>	4,8%	<b>100,0%</b>	85,0%
	Sim	0,0%	<b>95,2%</b>	0,0%	15,0%
Banners	Não	84,5%	100,0%	94,4%	91,8%
	Sim	<b>15,5%</b>	0,0%	5,6%	8,3%
SMS	Não	66,2%	<b>100,0%</b>	73,3%	75,0%
	Sim	33,8%	0,0%	26,7%	25,0%
Publicidade no Facebook	Não	41,5%	<b>100,0%</b>	73,8%	66,5%
	Sim	<b>58,5%</b>	0,0%	26,2%	33,5%
Televisão/Outros	Não	99,3%	95,2%	100,0%	99,0%
	Sim	0,7%	<b>4,8%</b>	0,0%	1,0%
Sexo	Feminino	59,2%	46,0%	61,0%	58,0%
	Masculino	40,8%	54,0%	39,0%	42,0%
Total		48 %	16 %	36 %	100%

Nota: A negrito identificaram-se as principais características dos grupo.

**Análise discriminante:**

De entre as variáveis selecionadas com base no valor absoluto valores da matriz de estrutura (anexo 10), verifica-se pela estatística Lambda de Wilks que os consumidores estão bem informados relativamente aos meios de publicidade, uma vez que a variável “não sei” não foi selecionada para o modelo (Tabela 61)

**Tabela 61:** Modelo final de Análise Discriminante obtido para os meios de publicidade mais eficazes, variável de grupo “Costuma realizar compras *online*”

	Wilks' Lambda	F	df1	df2	Sig.
Anúncios no Google / Sapo / ...	.993	2.949	1	398	.087
Emails publicitários	.975	10.142	1	398	.002
Banners	.979	8.578	1	398	.004
SMS	.995	1.832	1	398	.177
Publicidade no Facebook	.980	7.929	1	398	.005
Televisão/Outros	.996	1.550	1	398	.214

**Tabela 62:** Coeficientes da função de Fisher associada à Análise Discriminante obtido para os meios de publicidade mais eficazes, variável de grupo “Costuma realizar compras *online*”

	Costuma realizar compras Online?	
	Não	Sim
Anúncios no Google / Sapo / ...	14.902	14.478
Emails publicitários	13.607	14.303
Banners	20.553	21.826
SMS	15.142	14.638
Publicidade no Facebook	13.488	12.891
Televisão/Outros	121.618	123.135
(Constant)	-110.561	-112.458

Deste modo as funções de Fisher resultantes da análise discriminante são da forma:

$$D_{NComp}(x) = 14,902. \textit{anuncios\_google\_sapo} + 13,607. \textit{emails\_publicitarios} + 20,553. \textit{banners} \\ + 15,142. \textit{sms} + 13,488. \textit{Pub\_Facebook} + 121,618. \textit{televisão/Outros} - 110,561$$

$$D_{Comp}(x) = 14,478. \textit{anuncios\_google\_sapo} + 14,303. \textit{emails\_publicitarios} + 21,826. \textit{banners} \\ + 14,638. \textit{sms} + 12,891. \textit{Pub\_Facebook} + 123,135. \textit{televisão/Outros} - 112,458$$

Assim e considerando as pequenas diferenças entre os coeficientes de cada uma das funções de Fisher temos que não existem indícios de diferenças relevantes entre os meios de publicidade mais eficazes identificados pelos compradores e não compradores, confirmando as orientações do teste de Wilks.

Calculou-se, ainda, através da classificação efetuada a probabilidade de um novo elemento da amostra pertencer a um dos dois grupos (compradores ou não compradores), sendo que, como é possível verificar, apenas 46,8% dos não compradores são bem classificados, sendo que 76,1% dos compradores são bem classificados através deste modelo. Deste modo, constata-se que este modelo apresenta deficiências na classificação correcta dos não compradores.

**Tabela 63:** Probabilidade de um novo elemento pertencer ao grupo dos compradores, através do modelo obtido para os meios de publicidade mais eficazes

		Costuma realizar compras Online?	Predicted Group Membership		Total
			Não	Sim	
Original	%	Não	46.8	53.2	100.0
		Sim	23.9	76.1	100.0

a. 68.0% das observações iniciais foram corretamente classificadas.

Utilizando um raciocínio análogo, utilizando a variável sexo como *variável de agrupamento*, e usando o mesmo critério para escolha das variáveis, obtiveram-se os resultados apresentados na Tabela 64.

**Tabela 64:** Modelo final de Análise Discriminante obtido para os meios de publicidade mais eficazes, variável de grupo o sexo

	Wilks' Lambda	F	df1	df2	Sig.
Emails publicitários	.996	1.428	1	398	.233
Banners	.978	9.144	1	398	.003
SMS	.983	6.701	1	398	.010
Publicidade no Facebook	.988	4.904	1	398	.027
Televisão/Outros	.995	1.805	1	398	.180
Não sei	.993	2.713	1	398	.100



**Tabela 65:** Coeficientes da função de Fisher associada à Análise Discriminante para os meios de publicidade mais eficazes, usando como Variável de grupo "Sexo"

	Sexo	
	Feminino	Masculino
Emails publicitários	28.522	28.296
Banners	36.585	37.549
SMS	25.267	24.637
Publicidade no Facebook	29.499	28.942
Televisão/Outros	141.963	143.036
Não sei	49.442	49.506
(Constant)	-177.262	-177.614

Assim, as funções de Fisher resultantes da análise discriminante são da forma:

$$D_{Fem}(x) = 28,522. \text{emails\_publicitários} + 36,585. \text{banners} + 25,267. \text{SMS} \\ + 29,499. \text{publicidade\_facebook} + 141,963. \text{televisao\_outros} + 49,442. \text{nao\_sei} \\ - 177,262$$

$$D_{Masc}(x) = 28,296. \text{emails\_publicitários} + 37,549. \text{banners} + 24,637. \text{SMS} \\ + 28,942. \text{publicidade\_facebook} + 143,036. \text{televisao\_outros} + 49,506. \text{nao\_sei} \\ - 177,614$$

Assim e considerando as pequenas diferenças entre os coeficientes de cada uma das funções de Fisher temos que não existem indícios de diferenças relevantes entre os meios de publicidade mais eficazes identificados pelos dois géneros, confirmando as orientações do teste de Wilks.

Calculou-se, ainda, através da classificação efetuada probabilidade de um novo elemento da amostra pertencer a um dos dois grupos (homens ou mulheres), sendo que, como é possível verificar, 61,6% dos elementos do sexo feminino são bem classificados, que 63,7% dos homens são bem classificados através deste modelo.

**Tabela 66:** Percentagem de homens e mulheres classificados corretamente, através do modelo obtido para os meios de publicidade mais eficazes

		Sexo	Predicted Group Membership		Total
			Feminino	Masculino	
Original	%	Feminino	61.6	38.4	100.0
		Masculino	36.3	63.7	100.0

a. 62.5% das observações iniciais foram corretamente classificadas.

### Regressão logística:

Para construir um modelo de regressão logística para a variável “Costuma realizar compras *online*” considerando os meios de publicidade como variáveis independentes, introduziram-se todos os meios de publicidade *online* em simultâneo e foi feita uma seleção manual das variáveis utilizando o índice  $\left| \frac{S.E.}{\beta} \right|$ , sendo que quanto menores forem os valores deste rácio, maior será a contribuição da variável para o modelo. Utilizando este raciocínio, chegou-se ao seguinte modelo apresentado na Tabela 67, do qual é possível verificar que os emails publicitários e os *banners* são fatores potenciador de ser comprador *online*.

**Tabela 67:** Modelo de Regressão Logística obtido para os meios de *publicidade mais eficazes*, usando como variável suplementar “Costuma realizar compras *online*”

	B	S.E.	$\left  \frac{S.E.}{\beta} \right $	Exp(B)
Anúncios no Google/Sapo	-.431	.245	0,568	.650
Emails publicitários	.647	.216	0,334	1.910
Banners	1.905	.532	0,279	6.717
SMS	-.510	.240	0,471	.600
Publicidade no Facebook	-.580	.220	0,379	.560

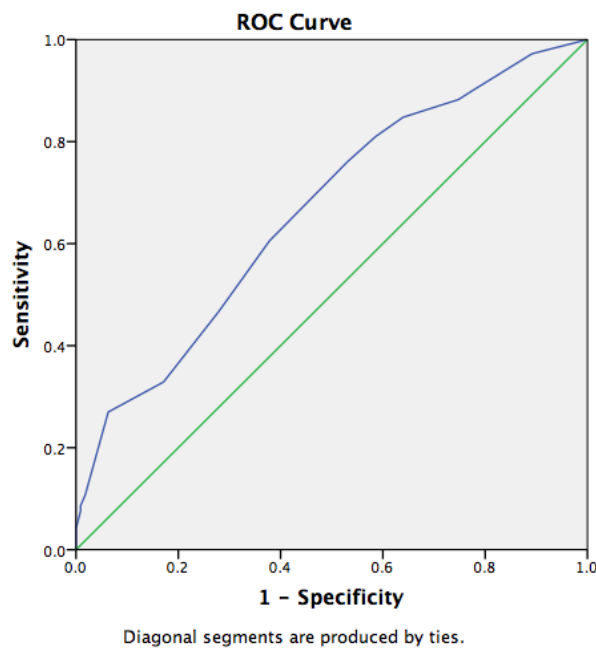
Para um ponto de corte de 0,70 obtemos um modelo com sensibilidade igual a 76,1% e uma especificidade igual a 46,8%, isto é são identificados corretamente 76,1% dos compradores e 46,8% dos não compradores, concluindo que o modelo ajusta corretamente 66,3% dos inquiridos o que representa uma qualidade razoável de ajustamento. O modelo encontrado tem ainda uma fraca capacidade discriminativa (AUC = 0,664)..

**Tabela 68:** Valores previstos com o Modelo de Regressão Logística para os compradores online em função dos meios de publicidade *online*, usando o ponto de corte 0,70

Observed		Predicted		
		Costuma realizar compras Online?		Percentage Correct
		Não	Sim	
Costuma realizar compras Online?	Não	52	59	46.8
	Sim	69	220	76.1
Overall Percentage				68.0

a. Ponto de corte = .700

**Gráfico 70:** Curva ROC associado ao Modelo de Regressão Logística obtido para os meios de publicidade mais eficazes, usando como variável suplementar “Costuma realizar compras online”



Através dos valores  $\hat{\beta}$ , podemos construir a função logística associada ao modelo obtido:

$$\text{logit}(\hat{\pi}(x)) = \beta_1 \cdot \text{anuncios\_no\_google\_sapo\_...} + \beta_2 \cdot \text{emails\_publicitarios} + \beta_3 \cdot \text{banners} + \beta_4 \cdot \text{SMS} + \beta_5 \cdot \text{publicidade\_no\_facebook}$$

$$= -0.431 \cdot \text{anuncios\_no\_google\_sapo\_...} + 0,647 \cdot \text{emails\_publicitarios} + 1,905 \cdot \text{banners} - 0,510 \cdot \text{SMS} - 0,580 \cdot \text{publicidade\_no\_facebook}$$

Utilizando um raciocínio semelhante, utilizou-se a regressão logitica para modelar o sexo em função dos meios de publicidade online identificados, tendo-se obtido o modelo presente na Tabela 69, que permitiu concluir que:

- Um indivíduo que prefira os Emails publicitários tem quase o dobro das possibilidades de ser mulher do que homem;
- Um indivíduo que prefira SMS tem mais do dobro das possibilidades de ser mulher do que homem;
- Um indivíduo que prefira Publicidade no Facebook tem quase o dobro das possibilidades de ser do mulher do que homem;

Repetindo o mesmo raciocínio, agora para modelar o sexo em função dos meios de publicidade mais eficazes, é possível verificar que os emails publicitários, as SMS e a publicidade no *Facebook* são fatores definidores dos elementos do sexo feminino. (Tabela 69).

**Tabela 69:** Modelo de Regressão Logística obtido para os meios publicitários mais eficazes, usando como variável suplementar “Costuma realizar compras *online*”

	B	S.E.	$\frac{ S.E. \beta }{\beta}$	Exp(B)
Emails publicitários	-.311	.209	0,672	.733
SMS	-.732	.251	0,343	.481
PublicidadenoFacebook	-.645	.228	0,353	.524
Constant	1.902	.588	0,309	6.697

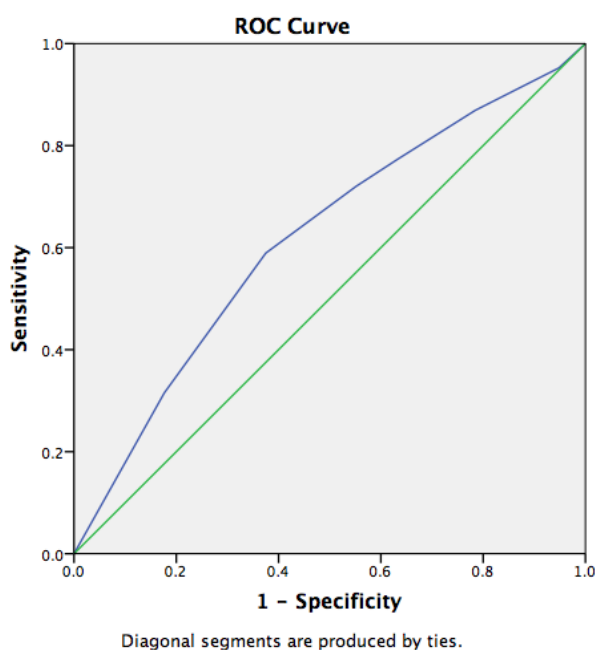
Assim, ao utilizarmos um ponto de corte de 0,40 obtemos um modelo com sensibilidade igual a 58,9% e uma especificidade igual a 62,5%, isto é, são identificados corretamente 58,9% dos indivíduos do sexo masculino e 62,5% dos indivíduos do sexo feminino, concluindo que o modelo ajusta corretamente 61,0% dos inquiridos o que representa uma boa qualidade de ajustamento. O modelo encontrado tem ainda uma fraca capacidade discriminativa (AUC = 0,618).

**Tabela 70:** Valores previstos com o Modelo de Regressão Logística para os sexos em função dos meios de publicidade *online*, usando o ponto de corte 0,40

Observed		Predicted		
		Sexo		Percentage Correct
		Feminino	Masculino	
Sexo	Feminino	145	87	62.5
	Masculino	69	99	58.9
Overall Percentage				61.0

a. Ponto de corte = .400

**Gráfico 71:** Curva ROC associado ao Modelo de Regressão Logística para o sexo em função dos meios de publicidade mais eficazes



Através dos valores  $\hat{\beta}$ , podemos construir a função logística associada ao modelo obtido:

$$\begin{aligned} \text{logit}(\hat{\pi}(x)) &= \beta_0 + \beta_1 \cdot \text{emails\_publicitarios} + \beta_2 \cdot \text{SMS} + \beta_3 \cdot \text{publicidade\_no\_facebook} \\ &= 1,092 - 0,311 \cdot \text{emails\_publicitarios} - 0,732 \cdot \text{SMS} - 0,645 \cdot \text{publicidade\_no\_facebook} \end{aligned}$$

Por via a comparar as variáveis e conclusões retiradas através da Análise de Correspondências múltipla, da função de Fisher resultante da Análise Discriminante e da Regressão Logística quando sujeitas ao mesmo problema, segue a tabela resumo dos meios de publicidade mais eficazes identificadas por compradores e não compradores e entre géneros entre os vários modelos utilizados (Tabela 71), concluindo-se que quer na Análise de Correspondências Múltiplas (ACM) quer através da Função de Fisher resultante da Análise Discriminante (AD) conclui-se que não existem diferenças entre compradores e não compradores e entre géneros no que diz respeito aos meios publicitários que mais impulsionam a compra *online* identificadas pelos indivíduos na amostra em estudo.

No que diz respeito às variáveis identificadas como relevantes na Análise Discriminante (AD) e na Regressão Logística (RL), registou-se que:

- Todas as variáveis escolhidas, para AD para os grupos construídos por homens/mulheres também foram selecionadas e por compradores/não compradores;
- Todas as variáveis escolhidas, para RL para os grupos construídos por homens/mulheres também foram selecionadas e por compradores/não compradores;
- Todas as variáveis selecionadas para a RL também foram selecionadas no modelo de AD quer para o grupo de compradores/não compradores quer para homens/mulheres;

**Tabela 71:** Tabela resumo dos métodos utilizados na análise dos meios de publicidade mais eficazes em comprar *Online*

	Compradores vs. Não Compradores		Homens vs. Mulheres	
	Existem indícios de diferenças entre os grupos?	Variáveis diferenciadoras	Existem indícios de diferenças entre os grupos?	Variáveis diferenciadoras
Análise de Correspondências Múltiplas	Não		Não	
Análise Discriminante -> Função de Fisher	Não		Não	
Regressão Logística	Sim	<ul style="list-style-type: none"> <li>• Banners</li> <li>• Emails publicitários</li> <li>• Anúncios no Google/Sapo/...</li> <li>• SMS</li> <li>• Publicidade no Facebook</li> </ul>	Sim	<ul style="list-style-type: none"> <li>• Emails publicitários</li> <li>• SMS</li> <li>• Publicidade no Facebook</li> </ul>

## 5) Considerações Finais

### 5.1) Limitações do estudo

Ao longo do estudo foram abordadas várias vezes as limitações inerentes ao processo de amostragem utilizado para a recolha da amostra, limitando este os métodos estatísticos a utilizar ao longo do estudo: a escolha de modos alternativos de seleção de variáveis, a não exploração de todo o potencial de cada um dos métodos (por via da não validação de tudo o que são pressupostos probabilísticos) assim como a validade das conclusões para um grupo de indivíduos distinto daquele em análise neste trabalho.

A comparação efetuada entre os vários métodos estudados merece igualmente validação em estudos futuros, de modo a suplantarmos eventuais enviesamentos por via da amostra. A obtenção de uma amostra representativa do universo Português é talvez o maior desafio neste tipo de estudos. Essa amostra poderia constituir uma excelente base para futuros trabalhos capazes de compreender de forma mais fidedigna o estado da arte do *e-commerce* em Portugal.

Apesar de abordadas ao longo do trabalho, será importante aprofundar mais as questões da publicidade *online*, de forma a compreender-se quais os meios de comunicação digital onde cada tipo de anunciante deverá investir o seu orçamento e quais as estratégias mais eficazes para cada um destes meios de comunicação. Estes tópicos são de facto importantes para um melhor trabalho na angariação de clientes e no crescimento do tecido empresarial português no meio *online*.

### 5.2) Conclusões sobre as hipóteses colocadas

De acordo com as análises efetuadas e os resultados obtidos (Capítulo 4) segue a tabela resumo das conclusões para cada uma das hipóteses colocadas (Secção 3.1) na construção deste estudo.

Note-se que todas as conclusões obtidas dizem apenas respeito essencialmente à amostra em estudo, constituindo assim uma base para estudos futuros que compreendam um planeamento amostral mais complexo, mais adaptado e com menos restrições do que o realizado neste estudo.



**Tabela 72:** Resumo dos resultados para as hipóteses do estudo

Hipótese	Conclusão
1) Não existem diferenças entre géneros no que diz respeito a comprar (ou não) online	<b>Hipótese validada</b>
2) Existem diferenças entre os produtos comprados por homens e por mulheres	<b>Hipótese validada</b>
3) Existem diferenças entre os serviços que homens e mulheres aderem online	<b>Hipótese não validada</b>
4) As inquiridas do sexo feminino tem maior tendência a comprar por impulso relativamente aos homens	<b>Hipótese não validada</b>
5) Existem diferenças entre as vantagens da compra <i>online</i> identificadas por compradores e não compradores	<b>Hipótese não validada</b>
6) Existem diferenças entre as vantagens da compra <i>online</i> identificadas por homens e mulheres	<b>Hipótese não validada</b>
7) Existem diferenças entre as desvantagens da compra <i>online</i> identificadas por compradores e não compradores	<b>Hipótese não validada</b>
8) Existem diferenças entre as desvantagens da compra <i>online</i> identificadas por homens e mulheres	<b>Hipótese não validada</b>
9) Existem diferenças entre os meios de comunicação digital preferidos por homens e mulheres	<b>Hipótese não validada</b>
10) Existem diferenças entre os meios de comunicação digital preferidos por compradores e não compradores	<b>Hipótese não validada</b>

Para além da validação destas hipóteses, outro dos objetivos deste trabalho passava pela tentativa de caracterizar os hábitos de consumo de *e-commerce*. Nesse sentido, e com base nas análises efetuadas, foi possível definirem-se vários perfis de consumo junto dos inquiridos. Um dos primeiros resultados obtidos foi a conclusão de que apenas existiam 5 tipos de compradores *online*: os generalistas, que compram todo o tipo de produtos discriminados neste trabalho, e os que compram outros produtos (não discriminados), os que compram produtos relacionados com atividades e experiências, os que compram produtos relacionados com moda e, finalmente, os que compram produtos relacionados com tecnologia e audiovisual, sendo este último grupo composto na sua maioria por homens. Concluiu-se também que quem compra brinquedos ou automóveis *online*, tem uma muito maior possibilidade de ser do sexo masculino do

que do sexo feminino.

No que concerne aos serviços aderidos *online*, a esmagadora maioria (95%) já tem experiência na adesão a todo o tipo de serviços na *internet*, tendo os homens uma maior tendência para usar este método de adesão. Curiosamente, apenas 1% desconhece que é possível aderir a estes serviços *online*.

Relativamente aos não compradores, concluiu-se que o principal motivo pelos quais estes não realizam compras *online* é não poder experimentar o produto ao vivo.

Identificaram-se também vários perfis de compradores relativamente às vantagens e às desvantagens identificadas por estes relativamente à compra *online*.

Finalmente estudaram-se também quais os meios de publicidade que os inquiridos julgavam ser mais eficazes, tendo-se identificado um perfil cujos membros consideram os *emails* de publicidade como o meio mais eficaz, outro perfil de indivíduos que valoriza mais anúncios televisivos ou outros meios publicitários e um último perfil de indivíduos que preferem meios de publicidade mais dirigidos aos seus interesses, como é o caso dos *banners* e da publicidade no *Facebook*.

No que diz respeito às comparações feitas entre a Análise de Correspondências Múltiplas (ACM), a Análise Discriminante (AD) e a Regressão Logística (RL), notou-se que, de um modo geral, existem algumas diferenças nas conclusões obtidas através dos vários métodos. Enquanto na ACM a procura de diferenças entre grupos concretos de indivíduos é feita de uma forma “passiva”, na AD e na RL todo o modelo é construído já nessa perspetiva. Apesar disso, as conclusões da ACM coincidem com as conclusões da AD, enquanto a RL se mostrou complementar a estes métodos ao ajudar na quantificação das diferenças entre os grupos em estudo.

## Bibliografia

- ACEPI (2010). Relatório Evolutivo do Barómetro do Comércio Electrónico em Portugal (1º Trimestre 2010).
- ANACOM (2004). *O Comércio Electrónico em Portugal – O Quadro Legal e o Negócio*, Lisboa: Autoridade Nacional de Comunicações.
- Benzécri, J. (1982). *Histoire et Préhistoire de L'Analyse des Données*. Paris: Dunod.
- Benzécri, J. (1984). *L'Analyse des Données (2 Vol.)*. Paris: Dunod.
- Bhanji, H. (2010). O Impacto da Aplicação do e-commerce num Website – O Caso da Samsung Electrónica Portuguesa. *Dissertação de Mestrado em Gestão*, ISCTE-IUL.
- Cabral, C. (2011). Publicidade na Internet – Análise Comparativa da Publicidade nos sites de três Jornais Online m Cabo Verde. Monografia apresentada para obtenção do grau de Licenciatura em Ciências da Comunicação, variante Publicidade. Universidade Jean Piaget de Cabo Verde.
- Cabrita, D. (2012). Métodos multivariados para variáveis qualitativas: aplicação ao estudo de variáveis associadas com a avaliação na disciplina de Matemática de uma escola do Ensino Básico no Concelho de Vila Nova de Gaia. *Tese de Mestrado em Estatística, Matemática e Computação, Especialidade em Estatística Computacional*. Universidade Aberta.
- Carvalho, H. (2008). *Análise Multivariada de Dados Qualitativos – Utilização da Análise de Correspondências Múltiplas com o SPSS*. Lisboa: Edições Sílabo.
- Castro, C. (2012). Comportamento do Consumidor: Procura de Informação e Intenção de Compra Online de Seguro Automóvel – Aplicação em Portugal. *Dissertação de Doutoramento em Gestão Empresarial Aplicada*, ISCTE Business School.
- Cheung, C., Zhu, L., Kwong, T., Chan, G., Limayem, M. (2003). Online Consumer Behavior: A Review and Agenda for Future Research. *16th Bled eCommerce Conference eTransformation*.
- Coleman, J. (1958). Relational analysis: The study of social organizations with survey methods. *Human Organization*, 17, 28–36.
- Comissão Europeia (2009). Europe's Digital Competitiveness Report. ICT Country Profiles. Vol. II.
- Conover, W. J. (1999). *Practical Nonparametric Statistics*. 3ª Edição. Wiley.
- Crivisqui, E. (1999). *PRESTA – Programme de Recherche et d'Enseignement en Statistique Appliquée*.
- Faugier, J., e Sargeant, M. (1997). Sampling hard to reach populations. *Journal of Advanced Nursing*, 26, 790–797.

- Ferreira, J. (2013). E-commerce e Género: Análise das Percepções de Compra do Consumidor. *Dissertação de Mestrado em Marketing*. Universidade de Lisboa, Instituto Superior de Economia e Gestão.
- Goodman, L. (1961). Snowball Sampling. *The Annals of Mathematical Statistics*, 32 (1), 148-170.
- Hill, M., e Hill, A. (2002). *Investigação por Questionário*. Lisboa: Edições Sílabo.
- Hoz, A. (1985). *Investigación Educativa: Dicionário Ciências da Educação*. Madrid: Ediciones Anaya, S.A.
- Lima, F. (2012). O Comércio Electrónico e as Plataformas B2C e C2C: contribuições para o estudo do comportamento do consumidor online. *Dissertação de Mestrado em Publicidade e Marketing*, Instituto Politécnico de Lisboa, Escola Superior de Comunicação Social.
- Lopes, A. (2014). Um estudo sobre as preferências dos consumidores em mercados digitais e tradicionais. *Dissertação de Mestrado em Marketing Digital*. Instituto Superior de Contabilidade e Administração do Porto, Instituto Politécnico do Porto.
- Maroco, J. (2007). *Análise Estatística com Utilização do SPSS*. Lisboa: Edições Sílabo.
- Nordeste, R. (2009). Publicidade online das empresas: estratégias actuais. Dissertação para obtenção do grau de Mestre em Comunicação Multimédia. Departamento de Comunicação e Arte, Universidade de Aveiro.
- Pestana, M., e Gageiro, J. (2008). *Análise de Dados para Ciências Sociais – A Complementaridade do SPSS*. Lisboa: Edições Sílabo.
- Ratner (2010), Variable selection methods in regression: Ignorable problem, outing notable solution, *Journal of Targeting, Measurement and Analysis for Marketing* **18**, 65–75
- Reis, E. (1997). *Estatística Multivariada Aplicada*. Lisboa: Edições Sílabo.
- Santos, N. (2011). Privacidade e o Comportamento do Consumidor Online – Um Modelo Explicativo da Intenção de Utilizar o Comércio Electrónico. *Tese submetida como requisito parcial para a obtenção do grau de Doutor em Gestão, Especialidade em Marketing*. Departamento de Marketing, Operações e Gestão Geral, ISCTE – IUL.
- Saraiva, C. (2012). Determinantes do Comportamento de Compra Online. *Dissertação de Mestrado em Publicidade e Marketing*, Instituto Politécnico de Lisboa, Escola Superior de Comunicação Social.
- Spreen, M. (1992). Rare populations, hidden populations and link-tracing designs—What and why?. *Bulletin Methodologie Sociologique*, 36, p. 34-58.
- Thomson (2009), The Importance of Structure Coefficients in Structural Equation Modeling Confirmatory Factor Analysis, Interdisciplinary Applications.

Turban, E., King, D., McKay, J., Marshall, P., Lee, J. e Viehland, D. (2008). *Electronic Commerce: A Managerial Perspective*. Upper Saddle River: Pearson.

Vicente, P., Reis, E., e Ferrão, F. (2001). *Sondagens – A amostragem como factor decisivo de qualidade*. 2ª Edição. Edições Sílabo.

Zhou, L.; Dai, L. Zhang. D. (2007). Online Shopping Acceptance Model: a critical survey of consumer factors in online shopping. *Journal of Electronic Commerce Research*, 8 (1), 41- 62.

# Anexos

## Anexo 1 – Questionário



UNIVERSIDADE  
DE ÉVORA

### Inquérito sobre Hábitos de e-Commerce

Este inquérito enquadra-se no âmbito do Mestrado em Modelação Estatística e Análise de Dados da Universidade de Évora e destina-se a compreender os hábitos de consumo sobre compras electrónicas em Portugal.

**\* Required**

**Idade \***

  
**Sexo \***  
**Distrito de residência \***  
**Estado Cívil \***  
**Número de filhos \***  
**Grau de Instrução \***  

### Experiência Online

**Quão à vontade se sente a usar a internet? \***

1 2 3 4 5 6 7 8 9 10

Nada à vontade           Totalmente à vontade

**Com que frequência usa a internet? \***

**Costuma realizar compras Online? \***

## Compras Online

### Que tipo de produtos já comprou online? \*

- Acessórios Pessoais (Relógios, Malas, Perfumes...)
- Automóvel ou outro veículo motorizado
- Bilhetes para Concertos/Espétaculos
- Brinquedos
- Experiências / Vouchers
- Filmes/Séries
- Livros
- Produtos para o Lar e Decoração
- Refeições
- Roupas
- Supermercado
- Tecnologia
- Viagens
- Other:

### Que tipo de serviços já contratou/subscreveu online? \*

- Entrega de compras em casa ou Refeições ao Domicílio
- Gás, Água ou Electricidade
- Mudança de casa
- Telecomunicações
- Nenhum, mas conheço esta possibilidade
- Nenhum, desconheço esta possibilidade
- Other:

### Costuma fazer compras... \*

- Para si
- Para oferecer

### Quantas vezes por ano faz compras online? \*

### Habitualmente, em que dias costuma comprar Online? \*

### Habitualmente, em que altura do dia costuma comprar online? \*

- Manhã (08:00-13:00)
- Tarde (13:00-19:00)
- Noite (19:00-00:00)
- Madrugada (00:00-08:00)
- É indiferente

### Costuma fazer compras Online... \*

- programadas.
- por impulso.

### Algum dos seguintes meios de publicidade já o motivaram a comprar online? \*

- Anúncios de texto no Google/Sapo/...
- Banners de publicidade
- Publicidade no Facebook
- Emails publicitários
- SMS
- Nenhum
- Não sei
- Other:



## Não compradores

### Porque motivos não compra Online? \*

(Escolha no máximo 2)

- Não estar em casa para receber o produto
- Não quer esperar pela entrega do produto
- Porque não sabe como comprar online
- Prefere experimentar o produto ao vivo
- Receio do produto poder ser estraviado
- Receio do produto ser cobrado e não receber o produto
- Receio do produto ser entregue danificado
- Receio que seja enviado do produto errado
- Other:

## Segurança nas compras Online

### Quão seguro considera ser comprar online? \*

1 2 3 4 5 6 7 8 9 10

Nada Seguro ○○○○○○○○○○ Totalmente Seguro

### Qual o factor que lhe transmite maior segurança relativamente a compras Online? \*

- A empresa/marca do produto
- O tipo de produto
- O site
- Os meios de pagamento disponíveis
- Other:

## Preferências

### Na sua opinião, qual/quais as principais vantagens de comprar online? \*

(Escolha no máximo 2)

- Não ter de me deslocar a uma loja
- Não ser influenciado por um vendedor
- Ter vantagens/promoções exclusivas
- Poder fazê-lo em qualquer hora/local
- Existência de um maior stock de produtos
- Possibilidade de fazer pré-reservas
- Poder ver e analisar comentários deixados por clientes anteriores
- Ter acesso às novidades
- Variedade de escolha
- Nenhuma
- Não sei
- Other:

### Na sua opinião, qual/quais as principais desvantagens de comprar online? \*

(Escolha no máximo 2)

- Dificuldades nas trocas/devoluções
- Envio do produto errado
- Método de pagamento
- Não ter apoio na escolha do produto
- Não experimentar o produto
- Portes de Envio
- Prazo de entrega
- Produto ser extraviado ou estar estragado
- Segurança
- Nenhuma
- Não sei
- Other:

**Para si, o que é que poderá servir de incentivo a comprar Online? \***

(Escolha no máximo 2)

- Portes de Envio Gratuitos
- Descontos Exclusivos Online
- Opção de entrega rápida
- Processo de compra simples e rápido
- Pressão de amigos/familiares
- Não sei
- Other:

**Qual tempo máximo que está disposto a esperar pela entrega de um produto que comprou online? \***

**Costuma comparar produtos online? \***

**Quais os meios de publicidade que considera mais eficazes para comunicar um desconto/promoção online? \***

(Escolha no máximo 2)

- Anúncios no Google / Sapo / ...
- Banners
- Publicidade no Facebook
- Emails publicitários
- SMS
- Não sei
- Other:

## Anexo 2 – Outputs complementares da ACM dos produtos comprados *online*

Tabela A.1: Medidas de discriminação para o modelo de ACM dos produtos comprados *online*

	Dimension														Mean
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
Bilhetes para Concertos/Espetáculos	.340	.147	.001	.015	.015	.048	.144	.007	.013	.073	.000	.014	.182	.001	.071
Acessórios Pessoais (Relógios)	.131	.211	.278	.000	.007	.000	.013	.004	.132	.034	.124	.013	.051	.001	.071
Automóvel ou outro veículo motorizado	.015	.008	.001	.080	.417	.282	.131	.004	.027	.028	.002	.000	.005	.001	.071
Brinquedos	.286	.048	.036	.109	.013	.046	.006	.001	.291	.027	.099	.025	.000	.013	.071
Livros	.066	.137	.272	.017	.019	.160	.088	.001	.026	.150	.020	.042	.000	.003	.071
Experiências / Vouchers	.292	.244	.093	.006	.010	.027	.024	.011	.006	.006	.011	.185	.007	.080	.071
Filmes/Séries	.150	.032	.293	.001	.014	.000	.049	.265	.056	.117	.004	.005	.001	.012	.071
Roupa	.101	.244	.231	.023	.003	.020	.034	.087	.031	.000	.186	.027	.004	.008	.071
Tecnologia	.202	.020	.033	.268	.060	.002	.000	.212	.037	.003	.136	.002	.025	.000	.071
Produtos para o Lar e Decoração	.376	.075	.023	.035	.008	.000	.125	.002	.004	.171	.003	.100	.056	.022	.071
Refeições	.339	.009	.001	.170	.026	.059	.183	.030	.018	.000	.001	.030	.012	.124	.071
Viagens	.323	.271	.000	.004	.010	.050	.012	.031	.016	.024	.003	.082	.121	.052	.071
Supermercado	.275	.003	.018	.274	.105	.025	.017	.071	.052	.008	.012	.011	.015	.113	.071
Outros	.016	.014	.000	.134	.386	.266	.036	.076	.019	.038	.010	.002	.002	.000	.071
Sexo*	.028	.009	.109	.012	.044	.021	.005	.050	.001	.016	.022	.000	.004	.001	.023
Active Total	2.911	1.464	1.280	1.135	1.093	.986	.863	.800	.729	.679	.610	.537	.483	.430	1.000
% of Variance	20.793	10.454	9.142	8.109	7.809	7.042	6.162	5.717	5.206	4.852	4.360	3.833	3.452	3.069	7.143

## Anexo 3 – Outputs complementares da ACM dos serviços aderidos *online*

 Tabela A.2: Medidas de discriminação para o modelo de ACM dos serviços *online*

	Telecomunicações	Entrega de compras em casa ou Refeições ao Domicilio	Gás, Água ou Electricidade	Mudança de casa	Nenhum, desconheço esta possibilidade	Sexo
Telecomunicações	1.000	.102	.377	.202	.101	.218
Entrega de compras em casa ou Refeições ao Domicilio	.102	1.000	.171	.187	.067	.007
Gás, Água ou Electricidade	.377	.171	1.000	.066	.081	.143
Mudança de casa	.202	.187	.066	1.000	.020	.092
Nenhum, desconheço esta possibilidade	.101	.067	.081	.020	1.000	.033
Sexo <sup>a</sup>	.218	.007	.143	.092	.033	1.000
Dimension	1	2	3	4	5	
Eigenvalue <sup>b</sup>	1.602	1.012	.959	.852	.575	

## Anexo 4 – Outputs complementares da AD dos serviços aderidos *online*

Tabela A.3: Modelo de AD do sexo em função dos serviços aderidos *online*

	Wilks' Lambda	F	df1	df2	Sig.
Telecomunicações	.953	14.300	1	287	.000
Entrega de compras em casa ou Refeições ao Domicilio	1.000	.013	1	287	.909
Gás, Água ou Electricidade	.979	6.024	1	287	.015
Mudança de casa	.991	2.476	1	287	.117
Nenhum, desconheço esta possibilidade	.999	.314	1	287	.575

Tabela A.4: Matriz de estrutura do modelo de AD do sexo em função dos serviços aderidos online

	Function
	1
Telecomunicações	.919
Gás, Água ou Electricidade	.597
Mudança de casa	.383
Nenhum, desconheço esta possibilidade	-.136
Entrega de compras em casa ou Refeições ao Domicilio	.028

## Anexo 5 – Outputs complementares da ACM das vantagens da compra *online*

Tabela A.5: Medidas de discriminação para o modelo de ACM das vantagens da compra *online*

Dimension	Cronbach's Alpha	Variance Accounted For		
		Total (Eigenvalue)	Inertia	% of Variance
1	.243	1.269	.159	15.868
2	.149	1.149	.144	14.369
3	.081	1.077	.135	13.457
4	.055	1.051	.131	13.137
5	-.019	.984	.123	12.299
6	-.031	.973	.122	12.165
7	-.259	.815	.102	10.193
8	-.536	.681	.085	8.511
Total		8.000	1.000	
Mean	.000 <sup>a</sup>	1.000	.125	12.500

a. Mean Cronbach's Alpha is based on the mean Eigenvalue.

## Anexo 6 – Outputs complementares da AD das vantagens da compra *online*

Tabela A.6: Modelo de AD dos compradores em função das vantagens da compra *online*

	Wilks' Lambda	F	df1	df2	Sig.
Não ser influenciado por um vendedor	.995	2.129	1	398	.145
Poder ver e analisar comentários deixados por clientes anteriores	.979	8.552	1	398	.004
Ter vantagens/promoções exclusivas	1.000	.010	1	398	.919
Variedade de escolha	.998	.906	1	398	.342
Outros	1.000	.015	1	398	.902
NenhumaNaosei	.996	1.550	1	398	.214
Comodidade	.995	1.943	1	398	.164
Novidades	.996	1.595	1	398	.207

Tabela A.7: Matriz de estrutura do modelo de AD dos compradores em função das vantagens da compra *online*

	Function
	1
Poder ver e analisar comentários deixados por clientes anteriores	-.734
Não ser influenciado por um vendedor	.366
Comodidade	.350
Novidades	.317
NenhumaNaosei	-.313
Variedade de escolha	.239
Outros	-.031
Ter vantagens/promoções exclusivas	-.026

Tabela A.8: Modelo de AD do sexo em função das vantagens da compra *online*

	Wilks' Lambda	F	df1	df2	Sig.
Não ser influenciado por um vendedor	1.000	.021	1	398	.885
Poder ver e analisar comentários deixados por clientes anteriores	.967	13.545	1	398	.000
Ter vantagens/promoções exclusivas	.996	1.418	1	398	.234
Variedade de escolha	1.000	.031	1	398	.861
Outros	.995	1.805	1	398	.180
Nenhuma Naosei	1.000	.106	1	398	.745
Comodidade	.995	1.886	1	398	.170
Novidades	.997	1.153	1	398	.284

Tabela A.9: Matriz de estrutura do modelo de AD do sexo em função das vantagens da compra *online*

	Function
	1
Poder ver e analisar comentários deixados por clientes anteriores	.820
Comodidade	-.306
Outros	.300
Ter vantagens/promoções exclusivas	-.265
Novidades	-.239
Nenhuma Naosei	.072
Variedade de escolha	.039
Não ser influenciado por um vendedor	.032



## Anexo 7 – Outputs complementares da ACM das desvantagens da compra *online*

Tabela A.10: Medidas de discriminação para o modelo de ACM das desvantagens da compra *online*

Dimension	Cronbach's Alpha	Variance Accounted For		
		Total (Eigenvalue)	Inertia	% of Variance
1	.285	1.339	.149	14.882
2	.264	1.306	.145	14.516
3	.183	1.195	.133	13.273
4	.168	1.175	.131	13.059
5	.073	1.069	.119	11.882
6	.049	1.046	.116	11.621
7	-.060	.949	.105	10.545
8	-.209	.843	.094	9.371
9	-13.564	.077	.009	.851
Total		9.000	1.000	
Mean	.000 <sup>a</sup>	1.000	.111	11.111

## Anexo 8 – Outputs complementares da AD das desvantagens da compra *online*

Tabela A.11: Modelo de AD dos compradores em função das desvantagens da compra *online*

	Wilks' Lambda	F	df1	df2	Sig.
Não experimentar o produto	.999	.552	1	398	.458
Portes de Envio	.961	16.275	1	398	.000
Produto ser extraviado ou estar estragado	1.000	.115	1	398	.735
Envio do produto errado	1.000	.087	1	398	.768
Método de pagamento	.993	2.867	1	398	.091
Segurança	.990	3.963	1	398	.047
Prazo de entrega	.975	10.142	1	398	.002
Falta de apoio no pré e pós-venda	.999	.481	1	398	.489
Nenhuma/Não sei	.997	1.278	1	398	.259

Tabela A.12: Matriz de estrutura do modelo de AD dos compradores em função das desvantagens da compra *online*

	Function
	1
Portes de Envio	.640
Prazo de entrega	.505
Segurança	-.316
Método de pagamento	-.269
Nenhuma/Não sei	-.179
Não experimentar o produto	-.118
Falta de apoio no pré e pós-venda	.110
Produto ser extraviado ou estar estragado	-.054
Envio do produto errado	.047

Tabela A.13: Modelo de AD do sexo em função das desvantagens da compra *online*

	Wilks' Lambda	F	df1	df2	Sig.
Não experimentar o produto	.995	1.865	1	398	.173
Portes de Envio	.999	.432	1	398	.511
Produto ser extraviado ou estar estragado	.978	9.066	1	398	.003
Envio do produto errado	1.000	.000	1	398	1.000
Método de pagamento	.994	2.307	1	398	.130
Segurança	.999	.260	1	398	.611
Prazo de entrega	.985	5.974	1	398	.015
Falta de apoio no pré e pós-venda	1.000	.069	1	398	.793
Nenhuma/Não sei	.996	1.477	1	398	.225

Tabela A.14: Matriz de estrutura do modelo de AD do sexo em função das desvantagens da compra *online*

	Function
	1
Produto ser extraviado ou estar estragado	.612
Prazo de entrega	.496
Método de pagamento	-.309
Não experimentar o produto	-.277
Nenhuma/Não sei	-.247
Portes de Envio	-.133
Segurança	.104
Falta de apoio no pré e pós-venda	.053
Envio do produto errado	.000

## Anexo 9 – Outputs complementares da ACM dos meios de publicidade mais eficazes

Tabela A.15: Medidas de discriminação para o modelo de ACM dos meios de publicidade mais eficazes

Dimension	Cronbach's Alpha	Variance Accounted For		
		Total (Eigenvalue)	Inertia	% of Variance
1	.400	1.523	.218	21.750
2	.324	1.385	.198	19.785
3	.090	1.084	.155	15.487
4	.021	1.018	.145	14.546
5	-.065	.947	.135	13.528
6	-.120	.907	.130	12.956
7	-7.393	.136	.019	1.947
Total		7.000	1.000	
Mean	.000 <sup>a</sup>	1.000	.143	14.286

a. Mean Cronbach's Alpha is based on the mean Eigenvalue.

## Anexo 10 – Outputs complementares da AD dos meios de publicidade mais eficazes

Tabela A.16: Modelo de AD dos compradores em função dos meios de publicidade mais eficazes

	Wilks' Lambda	F	df1	df2	Sig.
Anúncios no Google / Sapo / ...	.993	2.949	1	398	.087
Emails publicitários	.975	10.142	1	398	.002
Não sei	1.000	.177	1	398	.674
Banners	.979	8.578	1	398	.004
SMS	.995	1.832	1	398	.177
Publicidade no Facebook	.980	7.929	1	398	.005
Televisão/Outros	.996	1.550	1	398	.214

Tabela A.17: Matriz de estrutura do modelo de AD dos compradores em função dos meios de publicidade mais eficazes

	Function
	1
Emails publicitários	-.551
Banners	-.507
Publicidade no Facebook	.487
Anúncios no Google / Sapo / ...	.297
SMS	.234
Televisão/Outros	-.215
Não sei	.073

Tabela A.18: Modelo de AD do sexo em função das dos meios de publicidade mais eficazes

	Wilks' Lambda	F	df1	df2	Sig.
Anúncios no Google / Sapo / ...	.999	.537	1	398	.464
Emails publicitários	.996	1.428	1	398	.233
Banners	.978	9.144	1	398	.003
SMS	.983	6.701	1	398	.010
Publicidade no Facebook	.988	4.904	1	398	.027
Televisão/Outros	.995	1.805	1	398	.180
Não sei	.993	2.713	1	398	.100

Tabela A.19: Matriz de estrutura do modelo de AD do sexo em função dos meios de publicidade mais eficazes

	Function
	1
Banners	-.622
SMS	.532
Publicidade no Facebook	.455
Não sei	-.339
Televisão/Outros	-.276
Emails publicitários	.246
Anúncios no Google / Sapo / ...	-.151