

Instituto Superior de Contabilidade e Administração de Coimbra

Licenciatura de Informática de Gestão

Ano Letivo

2020 - 2021

Unidade Curricular de Sistemas Baseados em Conhecimento

Padrões dos Consumidores do Mercado Online de Entrega de Refeições ao Domicílio

Autor(es)

André Jesus Nº 2018071303

João Esculcas Nº 2018067460

Pedro Almeida Nº 2018069272

Elaborado em

08/02/2021

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

Índice

Índice.....	2
Índice de Figuras	3
Índice de Tabelas	4
Introdução.....	5
1. Entendimento do Negócio	6
1.1 Determinar os Objetivos de Negócio.....	6
1.2 Avaliação da Situação Atual.....	7
1.3 Definição dos Objetivos do Data Mining	7
1.4 Produzir o Plano do Projeto	8
2. Estudo dos dados	9
2.1 Recolha dos dados iniciais	9
2.2 Descrição dos dados	10
2.3 Exploração dos dados	13
2.4 Verificação da qualidade dos dados	23
3. Preparação dos dados.....	25
3.1 Seleção dos dados	25
3.2 Limpeza dos dados	25
3.3 Construção dos dados.....	26
3.4 Integração dos dados	27
3.5 Formatação dos dados.....	27
4. Modelação.....	29
4.1 Seleção da técnica de modelação	29
4.2 Geração do desenho de testes	30
4.3 Construção do modelo	31
4.4 Revisão do modelo.....	36
5. Avaliação	37
5.1 Avaliação de resultados	37
5.2 Revisão do processo	38
5.3 Determinar os próximos passos	38
6. Conclusão.....	39
Anexos	40
Referências.....	47

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

Índice de Figuras

Figura 1: Esquematização de Data Mining	5
Figura 2: Fases do modelo de referência CRISP-DM.....	8
Figura 3: Variáveis de Entrada.....	10
Figura 4: Sumário das Idades	13
Figura 5: Relação entre Idades e Regularidade de Encomendas de Refeições Online.....	13
Figura 6: Sumário do Género.....	14
Figura 7: Relação entre Género e Regularidade de Encomendas de Refeições Online	15
Figura 8: Sumário do Rendimento Mensal do agregado familiar	16
Figura 9: Relação entre Rendimento Mensal do Agregado Familiar e Regularidade de Encomendas de Refeições Online	16
Figura 10: Sumário do Nível de Escolaridade.....	17
Figura 11: Relação entre Nível de Escolaridade e Regularidade de Encomendas de Refeições Online.....	18
Figura 12: Sumário da Ocupação	19
Figura 13: Relação entre Ocupação e Regularidade de Encomendas de Refeições Online	19
Figura 14: Relação entre Estado Civil e Regularidade de Encomendas de Refeições Online ..	20
Figura 15: Relação entre Métodos de Encomenda e Regularidade de Encomendas de Refeições Online.....	21
Figura 16: Relação entre Motivos que Influenciam as Encomendas e Regularidade de Encomendas de Refeições Online	22
Figura 17: Relação entre Entraves às Encomendas e Regularidade de Encomendas de Refeições Online.....	22
Figura 18: Ficheiro CSV antes da formatação necessária	28
Figura 19: Ficheiro CSV depois da formatação necessária	28
Figura 20: Ficheiro CSV para dados de validação.....	30
Figura 21: Output do algoritmo SIMPLEKMEANS	31
Figura 22: Parâmetros utilizados	33
Figura 23: Outuput do modelo APRIORI.....	34

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

Índice de Tabelas

Tabela 1: Primeira parte do Inquérito.....	11
Tabela 2: Segunda parte do Inquérito.....	12
Tabela 3: Categorias da Idade.....	26
Tabela 4: Categorias do Rendimento mensal do agregado familiar	27
Tabela 5: Perfil do consumidor Típico.....	32
Tabela 6: Perfil obtido pelo cluster 0.....	32

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

Introdução

Este projeto foi desenvolvido no âmbito da unidade curricular de Sistemas Baseados em Conhecimento da licenciatura em Informática de Gestão do Instituto Superior de Contabilidade e Administração de Coimbra. A organização deste relatório está feita de acordo com a metodologia padrão CRISP-DM, um modelo de Data Mining, onde se pretende compreender o negócio em questão, entender e preparar os dados, de forma a permitir a aplicação de várias técnicas de modelagem que serão posteriormente avaliadas, sendo passíveis de implementação.

Os modelos descritivos tornaram-se indispensáveis para as pequenas e grandes empresas por permitem identificarem padrões desconhecidos comuns existentes em um conjunto de dados.

Com a desenvolver da tecnologia, é possível utilizar e analisar grandes bases de dados de forma rápida e eficiente, com o intuito de encontrar estratégias úteis que serão utilizadas pelas empresas para se destacarem, ou, nos tempos de pandemia que vivemos hoje, sobreviverem no mercado.

O presente projeto aborda a análise dos diversos fatores que influenciam o mercado de restauração online, como o intuito de auxiliar as empresas do setor da restauração a lidar com a nova realidade. Para isso, foi desenvolvido um processo de levantamento de dados através de um inquérito elaborado pelos constituintes do grupo de trabalho, onde todos os participantes deste estudo conservaram o seu anonimato



Figura 1: Esquematização de Data Mining

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

1. Entendimento do Negócio

Esta fase inicial foca-se no entendimento dos Objetivos e requisitos do projeto na perspetiva do negócio, convertendo então este conhecimento numa definição de problema de mineração de dados e na elaboração dum plano preliminar que permita alcançar os Objetivos definidos.

1.1 Determinar os Objetivos de Negócio

Devido à pandemia, o setor da restauração foi afetado em níveis nunca antes vistos. É nas palavras do primeiro-ministro, o setor “mais afetado” pelas medidas de restrição impostas pelo Governo.

Sem clientes e com encargos fixos, como o pagamento das rendas e dos salários, os restaurantes necessitam urgentemente de soluções que permitem manter parte da sua atividade e com isso ajudar a reduzir os enormes impactos negativos desta pandemia. Na nossa ótica, essa solução seria optar pelas novas plataformas de compra online de forma a aumentar as vendas.

A dependência dos smartphones não só veio revolucionar o nosso dia a dia, como também o mercado de entregas de refeições ao domicílio. Agora é comum pedir comida de quase qualquer tipo de restaurante, especialmente nas grandes cidades, sem precisar de comer “fast food”, e tudo através de uma aplicação. A forma tradicional de pedir comida, em que o consumidor contacta diretamente o restaurante para que lhe entregue o pedido em casa, está a ser substituída com o aparecimento das novas plataformas de compra online.

Há um notório desenvolvimento do ecossistema do comércio eletrónico em Portugal e a entrega de refeições ao domicílio aumentou, em 2018, 9% em relação a 2017. Este cenário reflete o impacto crescente dos serviços de entrega em Portugal como a Uber Eats e a Glovo.

De acordo com um estudo realizado, mais de 90% dos portugueses com acesso à internet já fizeram pelo menos uma compra online dentro de uma ampla gama de categorias. Os portugueses com acesso à internet compraram online especialmente viagens, produtos de moda, bilhetes para eventos, produtos de papelaria e produtos tecnológicos. No entanto, o dinamismo é claro no que diz respeito à compra de refeições para entrega ao domicílio.

Posto isto, é relevante compreender o comportamento do consumidor a fim de atender de forma ampla às suas necessidades, sendo este um dos objetivos primários do tema.

Objetivos primários:

- Manter os atuais clientes (na pior das hipóteses);
- Garantir a satisfação do cliente.

Objetivos secundários:

- Apresentar soluções que transmitam segurança e confiança a quem procura os seus serviços;
- Manter o índice de motivação de todos os seus colaboradores para que estes desenvolvam as suas funções com empenho e responsabilidade;
- Fomentar a satisfação de todos os clientes de modo a instituir vontade de regressar em futuras ocasiões e que estes recomendem os serviços aos seus familiares e amigos.

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

O principal critério de sucesso do tema é caso as receitas, provenientes das vendas, efetivamente aumentarem, critério este avaliado pela direção dos respetivos restaurantes.

1.2 Avaliação da Situação Atual

Intervenientes:

- Alunos:
 - André Jesus
 - João Esculcas
 - Pedro Almeida
- Docente:
 - Fernando Paulo Belfo

Dados trabalhados:

- Os dados com que trabalhamos neste trabalho provêm de um inquérito realizado na plataforma Google Forms. Estes dados foram recolhidos da população de Portugal continental.

Hardware:

- 3 Computadores Portáteis;

Software:

- Google Forms
- Excel
- Weka

1.3 Definição dos Objetivos do Data Mining

Data Mining é a principal etapa no processo de Descoberta de Conhecimento em Bases de Dados. É um conjunto de técnicas e ferramentas usadas para identificar padrões classificados de “conhecimento” embutidos em grandes massas de dados. É efetivamente a etapa de extração de padrões e/ou relacionamentos sistemáticos que tenham valor para o negócio.

Desta forma, com o Data Mining pretendemos perceber o motivo que leva aos consumidores optar pela entrega de refeições ao domicílio para se decidir que medidas adotar de modo a que se consiga maximizá-los ou, na pior das hipóteses, se mantenha o número e não haja reduções.

De acordo com estes objetivos desejados para a solução procurada, isto é, de acordo com o tipo de conhecimento que se espera extrair dos dados iremos recorrer a duas tarefas descritivas de mineração de dados: associação e agrupamento/segmentação/clustering.

Para que estes objetivos sejam concretizados com sucesso é necessário que os dados recolhidos sejam válidos e verídicos, pois, caso contrário o estudo não será legítimo. Para a questão da validade dos dados, decidimos elaborar o questionário com perguntas claras e diretas, de resposta fechada com campos formatados com o tipo de dados correto que cobrem todo o espectro de respostas possíveis/válidas de modo a evitar respostas fora do contexto da questão. Desta forma podemos garantir que não iremos receber respostas desnecessárias ao estudo. No entanto, concluímos que não dispomos de nenhuma forma de garantir a veracidade dos dados, confiando nas respostas dos inquiridos.

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

1.4 Produzir o Plano do Projeto

Para a correta realização deste trabalho foi necessário definir um plano do projeto, de forma a conseguir cumprir todas as normas e regras da metodologia CRISP-DM. Na metodologia em causa, o projeto é definido como um processo cíclico que permite alcançar as metas de mineração de dados e, por consequente, alcançar as metas de negócio. É relevante referir que a metodologia a seguir é descrita em termos dum modelo com processos hierárquicos, consistindo em conjuntos de tarefas descritas em quatro níveis de abstração (da generalidade à especificidade). Os níveis são as fases, as tarefas genéricas, as tarefas específicas e as instâncias do processo.

Embora o ciclo de vida de um projeto de mineração de dados consista em 6 fases, de acordo com o modelo de referência CRISP-DM, este plano do projeto divide-se apenas em cinco, pois não iremos implementar a última fase (desenvolvimento/instalação). Estas 5 fases estão divididas em 2 etapas distintas.

A primeira etapa consiste no estudo do projeto, abrangendo 3 fases. Na primeira é realizada uma análise ao negócio, percebendo a sua situação atual e como poderemos solucioná-la, delineando, assim, os objetivos a atingir com o projeto. Depois desta compreensão do projeto, segue-se o estudo dos dados (2ª fase), iniciando com a criação do conjunto de dados (obtidos através de um inquérito online) envolvendo as variáveis necessárias aos nossos propósitos. Após a recolha é feita uma descrição, exploração e verificação da qualidade dos dados. Ainda na primeira etapa é efetuada a preparação dos dados (3ª fase), onde é realizada a seleção dos dados de acordo com os critérios definidos, a limpeza, a reconfiguração, redução, combinação e transformação dos dados em formatos utilizáveis, disponibilizando-os, assim, de maneira usável e navegável para a realização do restante projeto.

A segunda etapa consiste na modelização e avaliação dos resultados, onde numa primeira fase será selecionada a técnica de modelação e construído o modelo e, numa segunda e última fase, será efetuada uma avaliação de resultados, os quais darão suporte à melhor tomada de decisões.

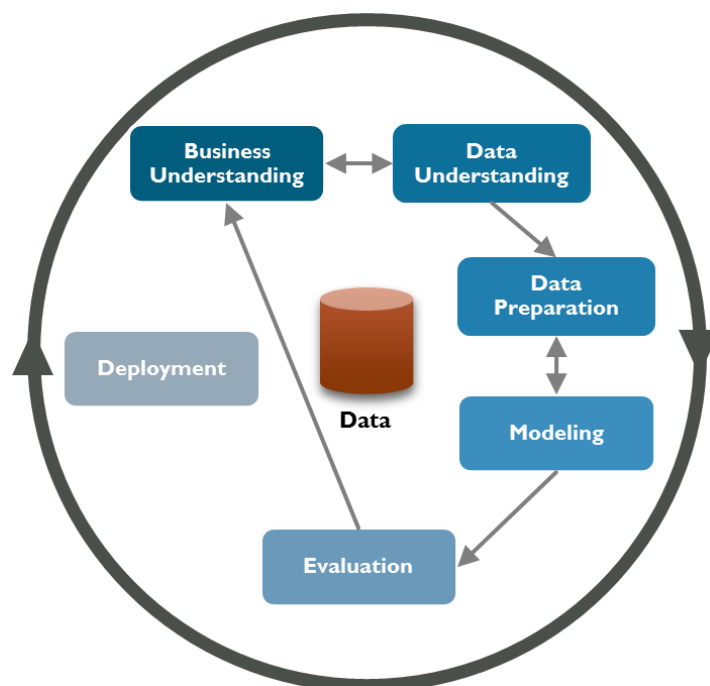


Figura 2: Fases do modelo de referência CRISP-DM

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

2. Estudo dos dados

A fase do estudo dos dados começa com a recolha dos dados iniciais a usar e desenvolve sobre eles atividades no sentido de se tornar cada vez mais familiar com eles, bem como no sentido de identificar alguns problemas que possam estar associados com a sua qualidade, descobrindo-lhes também alguns aspetos particulares e detetando subconjuntos interessantes de forma a poder formular hipóteses de informação desconhecida até ao momento.

2.1 Recolha dos dados iniciais

A tarefa de recolha de dados necessários ao nosso projeto consistiu num processo de criação de dados através de um questionário desenvolvido através da ferramenta Google Forms.

Após a estruturação do mesmo, o inquérito foi partilhado em diversas redes sociais de forma a obter o máximo número de respostas, e com o intuito de abranger o maior número de faixas etárias possíveis. No último dia disponível para a submissão do trabalho final, obtivemos um total de 114 respostas, no entanto, devido à antecipada data de entrega de uma primeira fase de estudo de dados, foram apenas contabilizados para análise um total de 113 respostas, das quais 90 já tinham encomendado uma refeição online.

Dada a necessidade de análise detalhada de dados, exportámos todos os dados obtidos para o Microsoft Excel e demos início ao estudo e preparação de dados.

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

2.2 Descrição dos dados

Em relação à descrição dos dados, devemos perceber e examinar as principais propriedades dos dados adquiridos. Estes dados foram recolhidos através do inquérito já previamente mencionado, tendo 7 perguntas base para todos os participantes e uma extensão de 4 perguntas para o público-alvo, que corresponde aos que já efetuaram alguma encomenda de uma refeição online.

Posto isto conseguimos resumir o estudo em questão através das seguintes variáveis de entrada.

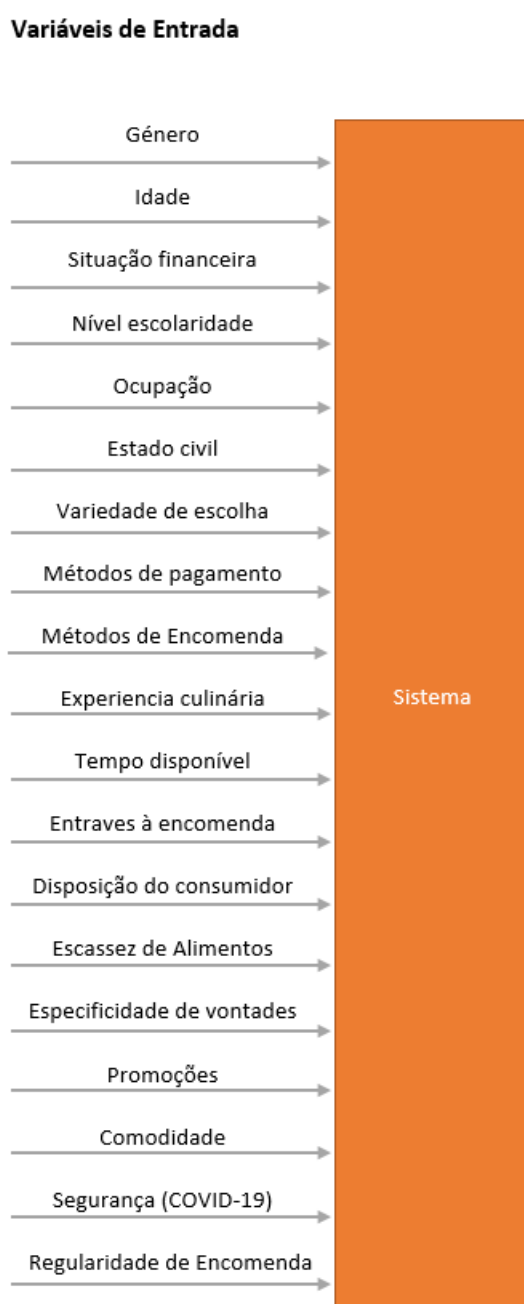


Figura 3: Variáveis de Entrada

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

De forma a obter uma melhor leitura e compreensão dos dados, agrupamos toda a informação exigida nas seguintes tabelas.

Atributo	Tipo/Formato	Domínio do Atributo	Objetivos Gerais
Idade	Nominal (Seleção)	<15; 15-20; 21-25; 26-30; > 30;	- “Perceber a faixa etária dos consumidores do mercado de restauração online”
Género	Nominal (Seleção)	Masculino; Feminino;	- “Perceber o género dos consumidores do mercado de restauração online”
Rendimento mensal do agregado familiar	Nominal (Seleção)	<600€; 600€ - 800€; 800€ - 1200€; 1200€ - 1600€; 1600€ - 2000€; > 2000€;	- “Entender como é que o rendimento familiar influencia o consumo”
Nível de Escolaridade	Nominal (Seleção)	-1º Ciclo; -2º Ciclo; -3º Ciclo; -Ensino Secundário; -Licenciatura; -Mestrado; -Doutoramento; -Outro;	- “Perceber o Nível de Escolaridade dos consumidores do mercado de restauração online”
Ocupação	Nominal (Seleção)	-Estudante; -Trabalhador; -Estatuto de trabalhador estudante; -Desempregado;	- “Perceber a Ocupação dos consumidores do mercado de restauração online”
Estado Civil	Nominal (Seleção)	-Solteiro; -Casado; -Divorciado; -Viúvo; -Outro;	- “Perceber o Estado Civil dos consumidores do mercado de restauração online”
Já alguma vez encomendou uma refeição online?	Nominal (Seleção)	Sim; Não;	- “Atributo que permite identificar a população alvo”

Tabela 1: Primeira parte do Inquérito

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

Atributo	Tipo/Formato	Domínio do Atributo	Objetivos Gerais
Regularidade encomenda mensalmente refeições online	Número Inteiro	Resposta livre	- “Atributo que permite fazer a relação entre os fatores de sucesso da restauração online”
Métodos usuais para encomendar	Nominal (Múltipla Seleção)	-Aplicações tipo uber eats, glovo, etc; -Aplicação da cadeia de restauração; -Contactar o restaurante através do contacto telefónico;	- “Perceber se a amostra se insere maioritariamente no mercado em questão”
Principais fatores que o levam a encomendar uma refeição	Nominal (Múltipla Seleção)	-Falta de tempo para cozinhar; -Pouca (ou nenhuma) experiência culinária Falta de vontade para cozinhar; -Escassez de alimentos em casa; -Vontade de um tipo específico de refeição; -Variedade de escolha; -Promoções; -Comodidade; -Diversidade de métodos de pagamento; -Segurança perante as novas circunstâncias (COVID-19); -Outro;	- “Motivos que levam os consumidores a optar pelo mercado de restauração online”
Entraves ao encomendar uma refeição online	Nominal (Múltipla Seleção)	-Preço da refeição; -Tempo de espera; -Carência do método de pagamento preferível; -Possíveis complicações associadas ao erro humano nos restaurantes; -Receio perante as novas circunstâncias (COVID-19); -Nenhum; -Outro;	- “Possíveis objeções dos consumidores à escolha pelo mercado da restauração online”

Tabela 2: Segunda parte do Inquérito

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

2.3 Exploração dos dados

O processo de exploração de dados é a base da análise e interpretação dos resultados obtidos em inquérito, de forma a elaborar uma conclusão relativamente ao nosso projeto. Com este objetivo em mente, decidimos comparar individualmente as variáveis de entrada com a “Regularidade de encomenda”, por ser o atributo que melhor reflete a influencia dos fatores.

Esta comparação é observável nos gráficos que se seguem.

2.3.1 Relação entre Idades e Regularidade de Encomendas de Refeições Online

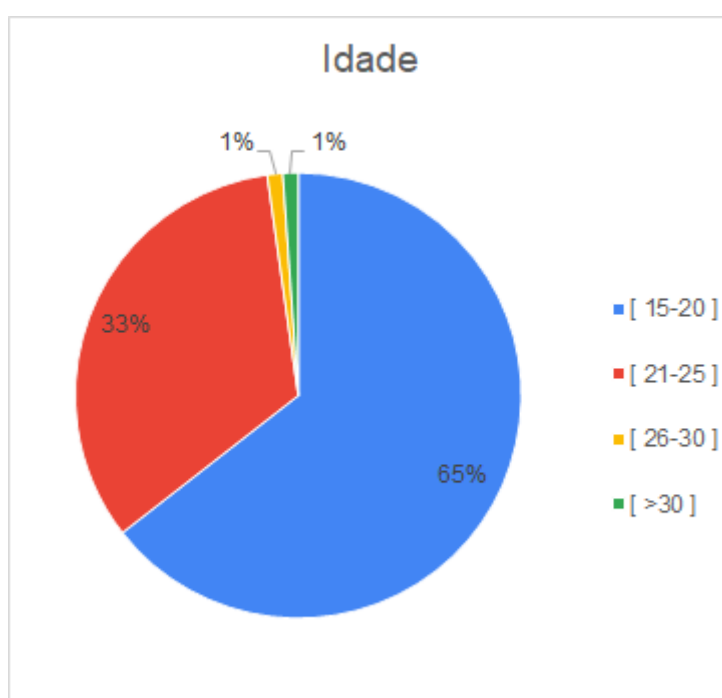


Figura 4: Sumário das Idades

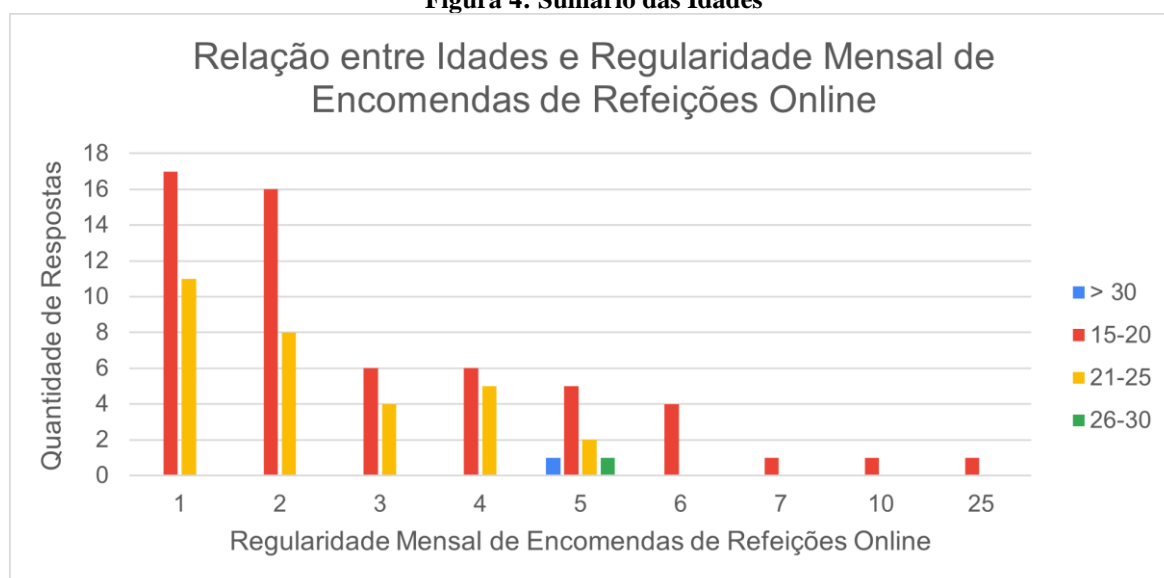


Figura 5: Relação entre Idades e Regularidade de Encomendas de Refeições Online

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

A ilustração 1 mostra a idade dos inquiridos. Entre as 90 respostas válidas observa-se que a grande maioria se encontra entre os 15 aos 20 anos (65%), seguida de 33% na faixa etária dos 21 aos 25 anos. Assim, apesar da idade dos inquiridos variar dos 15 aos 30 anos, a faixa etária predominante é a dos 15 aos 20 anos.

A ilustração 2 representa a relação das idades com a regularidade de encomendas de refeições online por parte da nossa amostra composta por 90 consumidores. Esta relação tem como objetivo dar-nos a entender se a idade tem influência na regularidade de encomendas de refeições online. Ao analisar o gráfico constatámos que a maior regularidade se encontra presente nos inquiridos de idade entre os 15 e os 25 anos. No entanto, como a nossa amostra de consumidores com faixa etária diferente à predominante é pouco significativa, não podemos concluir se a constatada predominância das idades influencia a regularidade de encomendas de refeições online.

2.3.2 Relação entre Género e Regularidade de Encomendas de Refeições Online

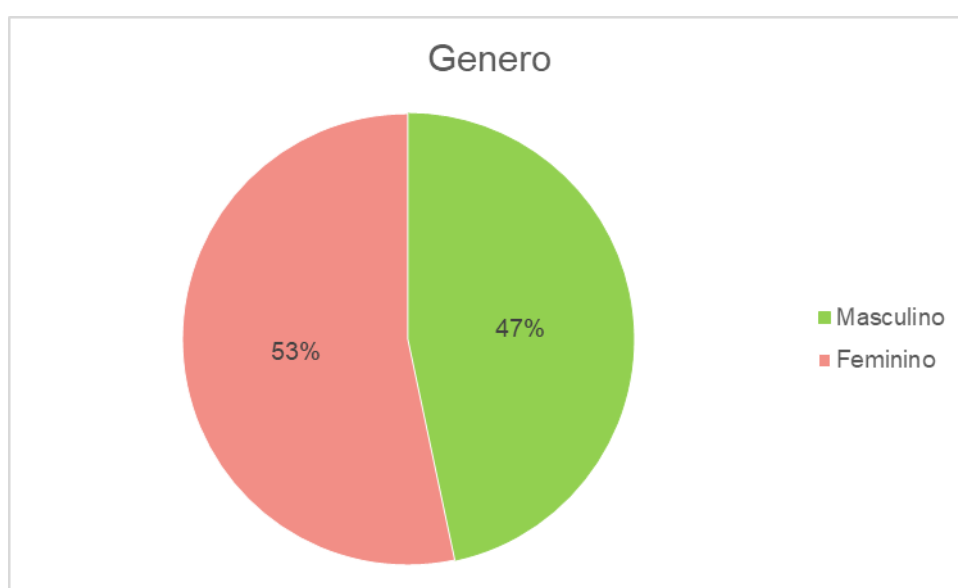


Figura 6: Sumário do Género

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

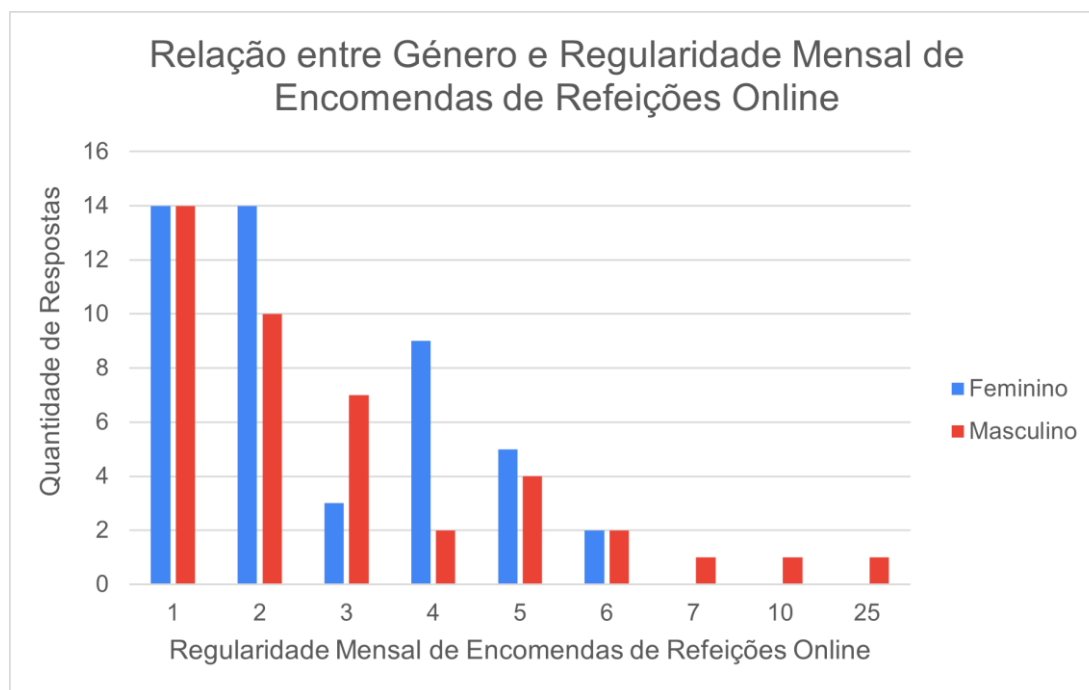


Figura 7: Relação entre Género e Regularidade de Encomendas de Refeições Online

A ilustração 1 mostra o género dos inqueridos. Entre as 90 respostas validas, observa-se que ambos os géneros são representados neste estudo sendo que 53% dos participantes são do sexo feminino e os restantes 47% são do sexo masculino.

A ilustração 2 representa a relação entre o género e a regularidade de encomendas de refeições online mensalmente, por parte da nossa amostra composta por 90 consumidores. Esta relação tem como objetivo dar-nos a entender se o género tem influência na regularidade de encomendas de refeições online. Ao analisar o gráfico constatamos que a maior regularidade se encontra em pessoas do sexo masculino, sendo o máximo de encomendas de refeições 25 mensalmente. No entanto os participantes do sexo feminino não fazem encomendas com tanta regularidade, mas são clientes mais certos, isto é, encomendam menos vezes, mas são mais pessoas do sexo feminino a encomendar. Posto isto, de acordo com os dados que temos, os participantes do sexo masculino fazem cerca de 261 encomendas de refeições online mensalmente sendo apenas 42 elementos, enquanto os participantes do sexo feminino fazem cerca de 233 encomendas de refeições online mensalmente tendo 48 elementos.

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

2.3.3 Relação entre Rendimento Mensal do Agregado Familiar e Regularidade de Encomendas de Refeições Online

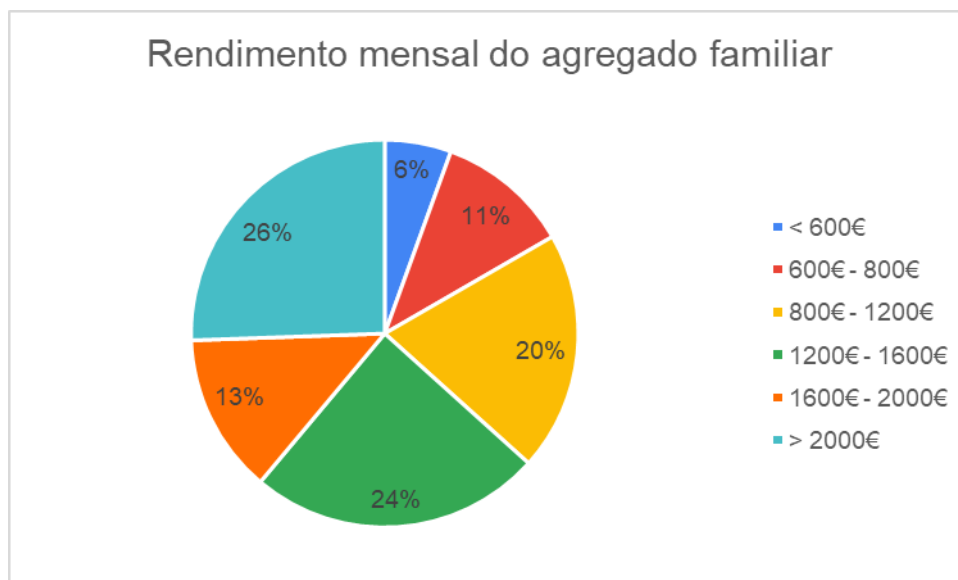


Figura 8: Sumário do Rendimento Mensal do agregado familiar

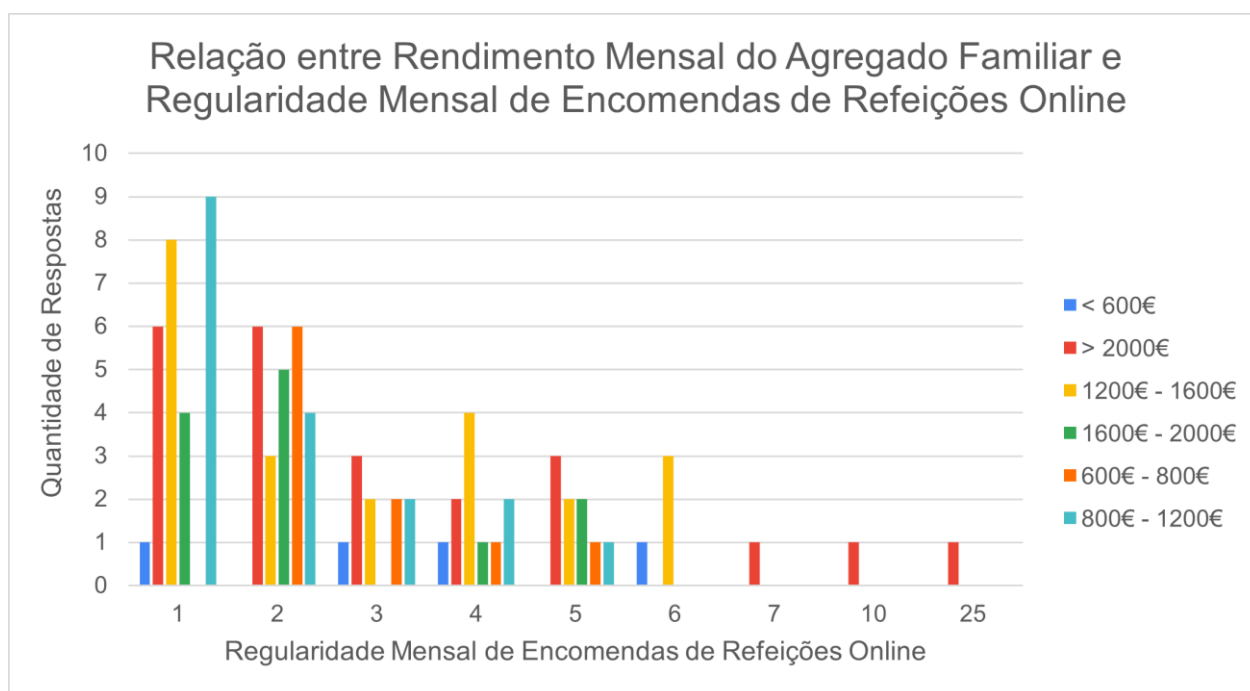


Figura 9: Relação entre Rendimento Mensal do Agregado Familiar e Regularidade de Encomendas de Refeições Online

A ilustração 1 mostra o rendimento mensal do agregado familiar de cada um dos inquiridos. Entre as 90 respostas validas observa-se que a maioria de pessoas (26%) tem um rendimento superior a 2000€, 24% com um rendimento entre 1200€ e 1600€, 20% com um rendimento entre 800€ e 1200€, 13% com um rendimento entre 1600€ e 2000€, 11% com um rendimento entre 600€ e 800€ e apenas 6% com um rendimento inferior a 600€. Podemos concluir que 94% está acima do salário minino português (700€), tendo em conta que o agregado familiar pode ser composto por apenas 1 elemento, é seguro assumir que todos os inquiridos têm capacidade de suportar um gasto não essencial como este.

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

A ilustração 2 representa a relação entre o rendimento mensal do agregado familiar e regularidade de encomendas de refeições online por parte da nossa amostra composta por 90 consumidores. Esta relação tem como objetivo dar-nos a entender se rendimento mensal do agregado familiar tem influência na regularidade de encomendas de refeições online. Ao analisar o gráfico constatámos que os inquiridos com um rendimento superior a 2000€ são os que fazem mais encomendas de refeições online com cerca de 92 encomendas mensais, no entanto cada inquerido deste grupo tem tendência a fazer a ter um maior volume de encomendas por mês. Como era de esperar o grupo com rendimento inferior a 600€ encomenda apenas 14 vezes sendo o grupo que tem o menos volume de encomendas por mês. Podemos concluir então que o rendimento mensal do agregado familiar não tem um peso muito significativo no processo de decisão de fazer uma encomenda da refeição online.

2.3.4 Relação entre Nível de Escolaridade e Regularidade de Encomendas de Refeições Online

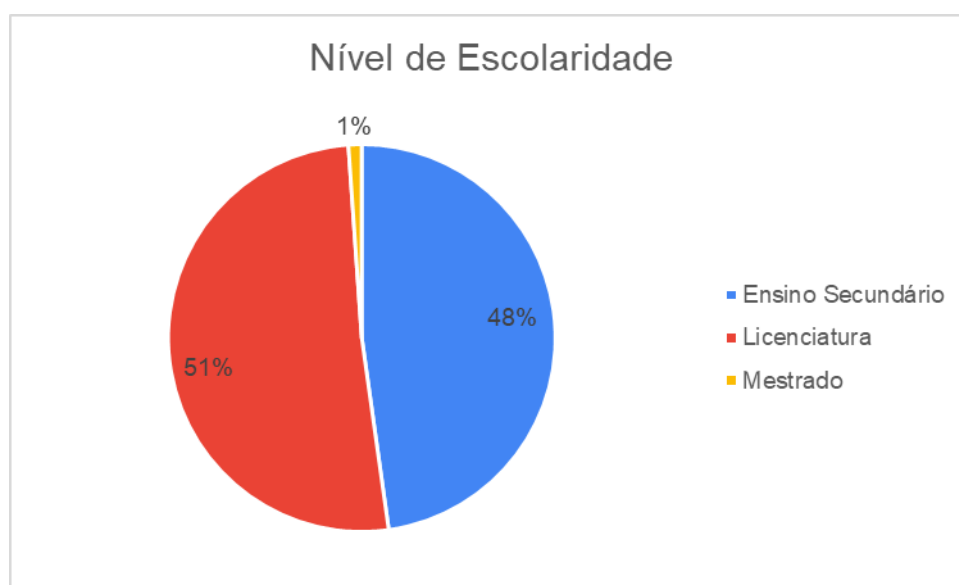


Figura 10: Sumário do Nível de Escolaridade

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

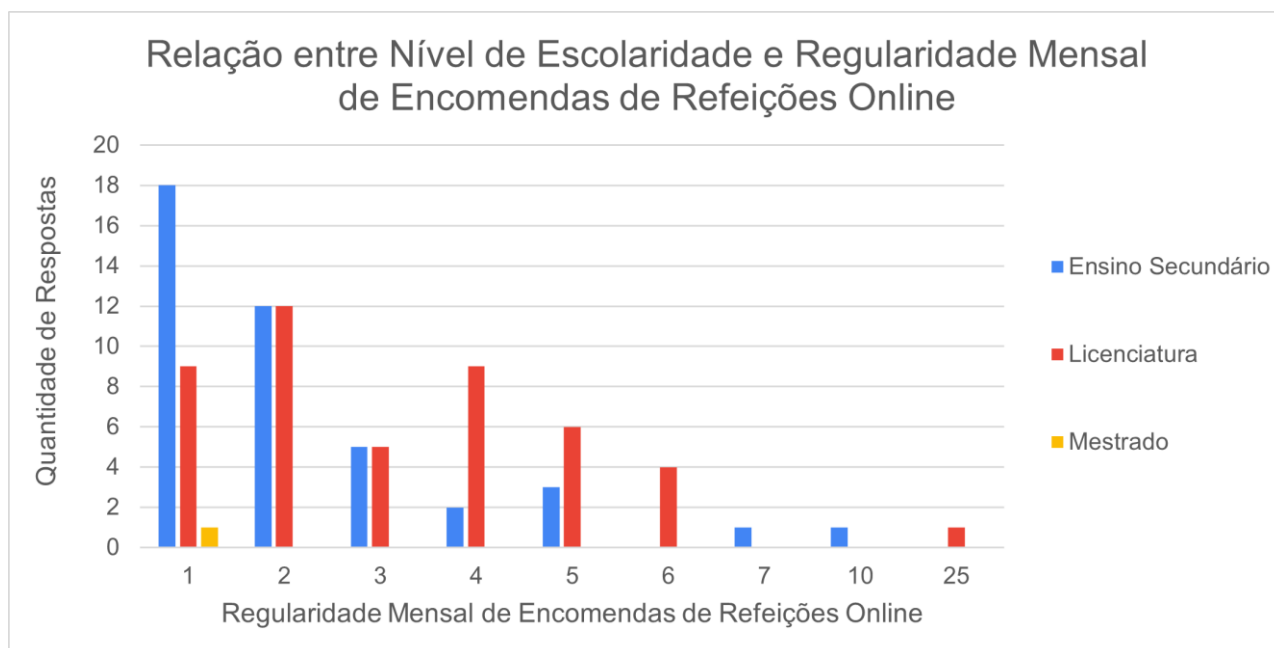


Figura 11: Relação entre Nível de Escolaridade e Regularidade de Encomendas de Refeições Online

A Ilustração 1 mostra o nível de escolaridade dos inqueridos. Entre as 90 respostas válidas observa-se que a maioria se encontra com a licenciatura feita (51%), seguido pelo ensino secundário (48%) e por fim mestrado (1%). Com estes dados podemos concluir que apenas estudantes tanto universitários como do ensino secundário encomendam refeições online.

A Ilustração 2 relação entre nível de escolaridade e regularidade de encomendas de refeições online por parte da nossa amostra composta por 90 consumidores. Esta relação tem como objetivo dar-nos a entender se o nível de escolaridade tem influência na regularidade de encomendas de refeições online. Ao analisar o gráfico constatamos que os inqueridos pertencentes ao grupo do ensino secundário constituem cerca de 37% do volume de encomendas de refeições online mesmo sendo cerca de metade de todos os participantes no estudo o que nos leva a concluir que o grupo dos inqueridos com licenciatura, domina com 63% do volume de encomendas de refeições online, apesar de ter praticamente o mesmo número de elementos que o grupo do ensino secundário. Como não temos mais níveis de ensino para comparar, não podemos fazer uma conclusão correta, no entanto podemos concluir que os participantes com licenciatura, encomendam refeições online com bastante regularidade.

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

2.3.5 Relação entre Ocupação e Regularidade de Encomendas de Refeições Online

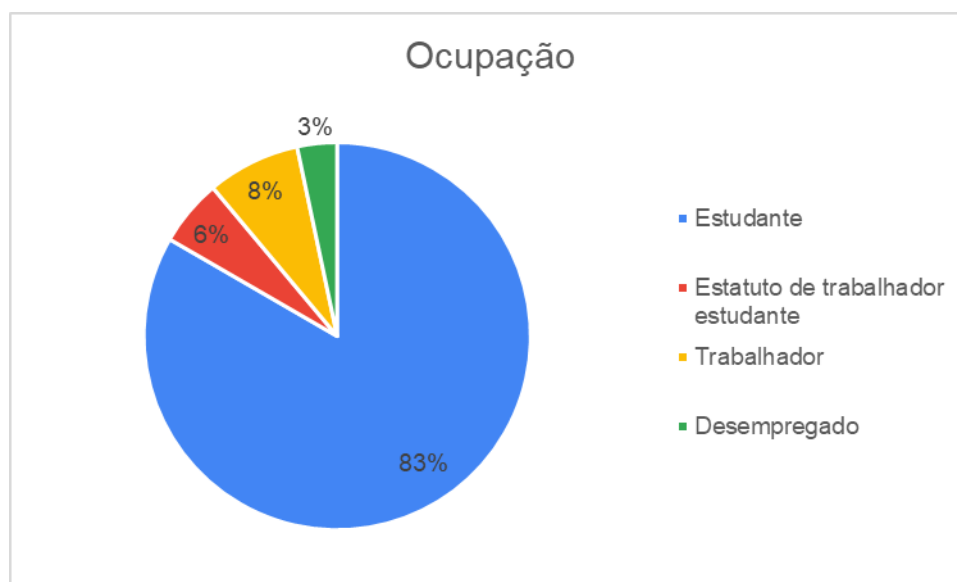


Figura 12: Sumário da Ocupação

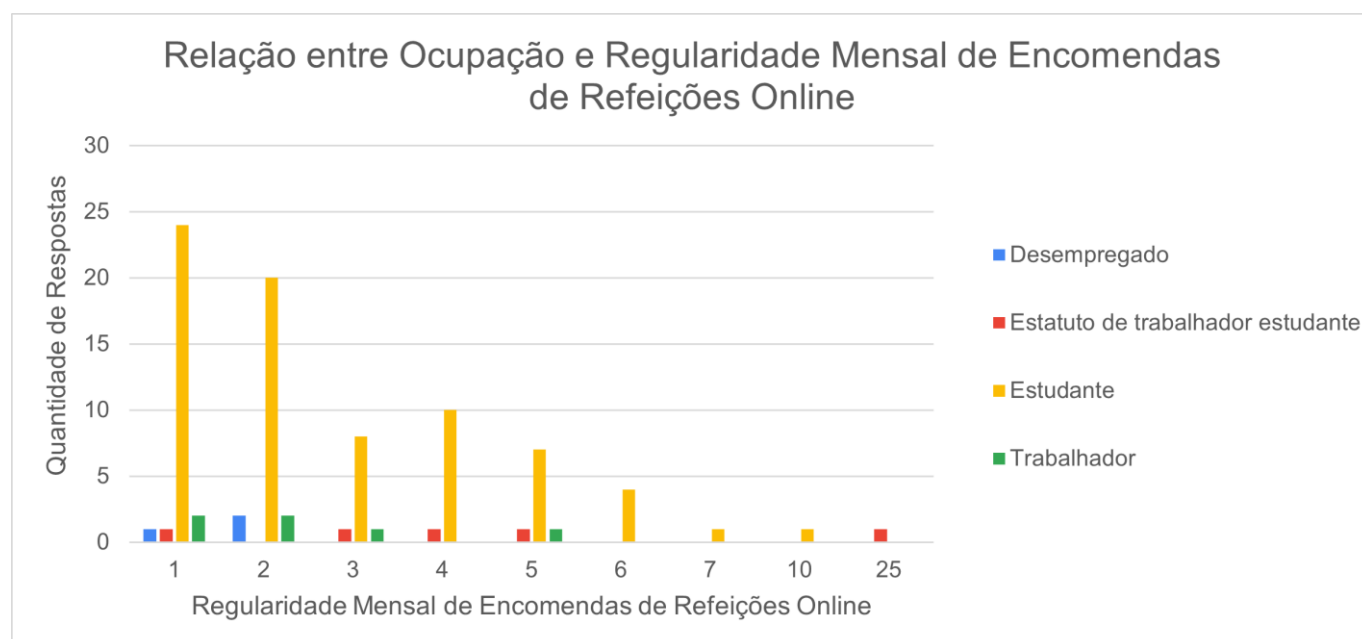


Figura 13: Relação entre Ocupação e Regularidade de Encomendas de Refeições Online

A ilustração 1 mostra a ocupação dos inqueridos. Entre as 90 respostas valias observa-se que a grande maioria são estudantes (83%), 8% são trabalhadores, 6% têm o estatuto de trabalhador estudante e apenas 3% estão desempregados.

A ilustração 2 representa a Relação entre ocupação e regularidade de encomendas de refeições online por parte da nossa amostra composta por 90 consumidores. Esta relação tem como objetivo dar-nos a entender se a ocupação tem influência na regularidade de encomendas de refeições online. Ao analisar o gráfico constamos que a principal ocupação da amostra é ser estudante.

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

2.3.6 Relação entre Estado Civil e Regularidade de Encomendas de Refeições Online

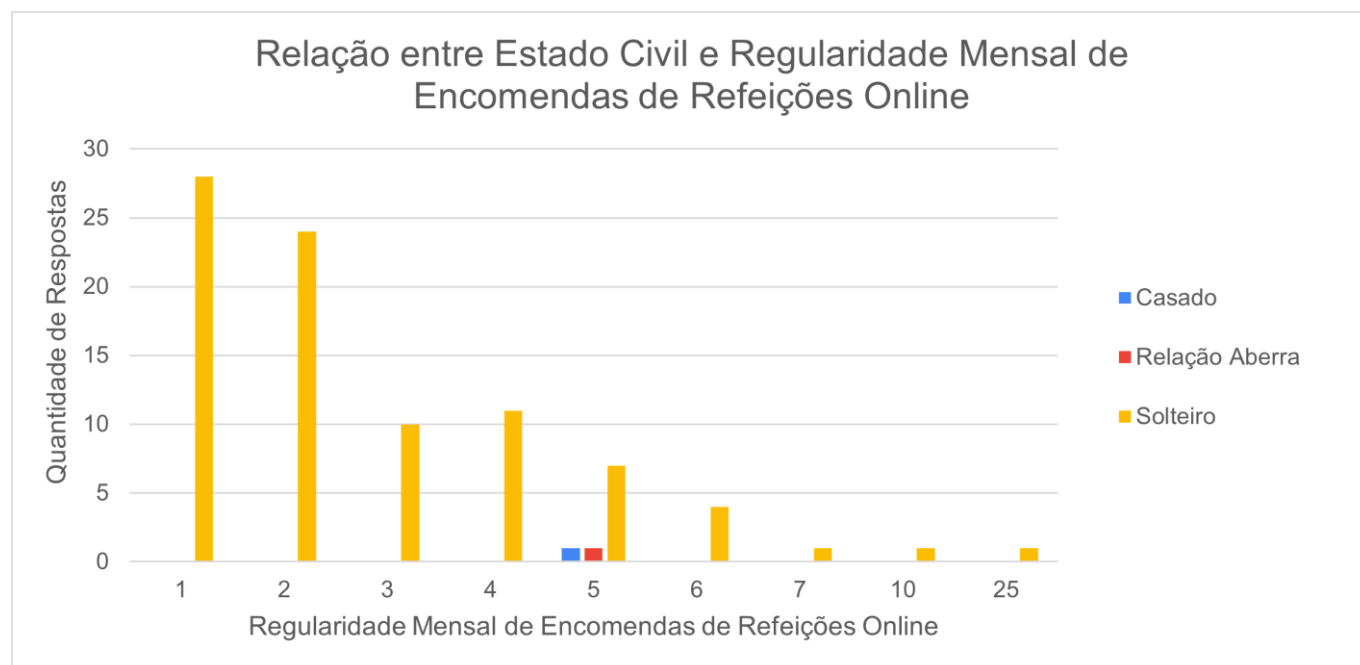


Figura 14: Relação entre Estado Civil e Regularidade de Encomendas de Refeições Online

Perante as 90 respostas válidas, constatamos que a grande maioria dos constituintes do estudo e simultaneamente integrantes do mercado de restauração online, encontram-se solteiros, constituindo 98% da amostra. Os restantes 2% subdividem-se em 2 casados 1 integrante da resposta aberta “Relação Aberra” futuramente devidamente filtrada.

Relativamente à ilustração, esta tem como objetivo verificar uma possível influência do estado civil na regularidade de encomendas de refeições online.

Ao analisar o gráfico, conseguimos evidentemente perceber que o estado civil predominante na amostra de mercado de restauração online presente no nosso estudo é o estado solteiro.

No contexto da nossa amostra, podemos então observar que o estado civil solteiro é mais propício a encomendar refeições online.

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

2.3.7 Relação entre Métodos de Encomenda e Regularidade de Encomendas de Refeições Online

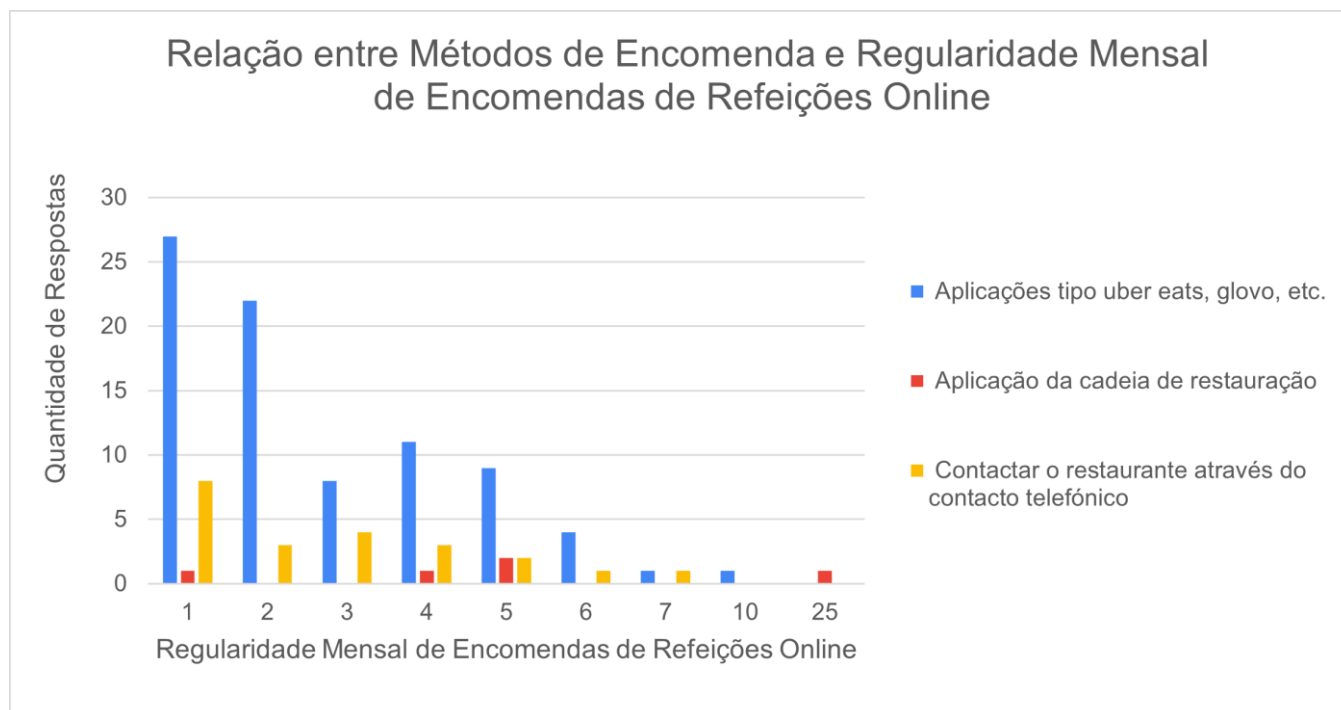


Figura 15: Relação entre Métodos de Encomenda e Regularidade de Encomendas de Refeições Online

De entre as 90 respostas consideradas no estudo, a amostra divide-se em 3 diferentes categorias no que toca aos métodos de encomenda: Aplicações como a Uber Eats ou a Glovo, aplicações da cadeia de restauração de que pretendem encomendar, ou através do contacto telefónico do restaurante. Uma vez que a questão permitia selecionar mais do que uma opção, observámos que 92,3% da população (correspondente a 84 inquiridos) utiliza aplicações como a Uber Eats ou Glovo, 24,2% (correspondente a 22 inquiridos) contacta o restaurante diretamente através do telefone, e 5,5% (correspondente a 5 inquiridos) utiliza a aplicação da cadeia de restauração da qual pretende encomendar.

A ilustração acima, mostra-nos a influência do método de encomenda, na regularidade com que a efetua. Ao analisar o gráfico, observamos imediatamente a preferência pelas aplicações como a Uber Eats ou Glovo. No entanto, após uma análise mais detalhada, apercebemo-nos que, independentemente da regularidade mensal de encomendas de refeições online, todos os métodos de encomenda são utilizados, inclusivamente, o inquirido da amostra que mais encomenda, serve-se da aplicação da cadeia de restauração para encomendar refeições.

No contexto da nossa amostra, podemos então observar que o método de encomenda não tem influência na regularidade de encomenda de refeições online.

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

2.3.8 Relação entre Motivos que Influenciam as Encomendas e Regularidade de Encomendas de Refeições Online

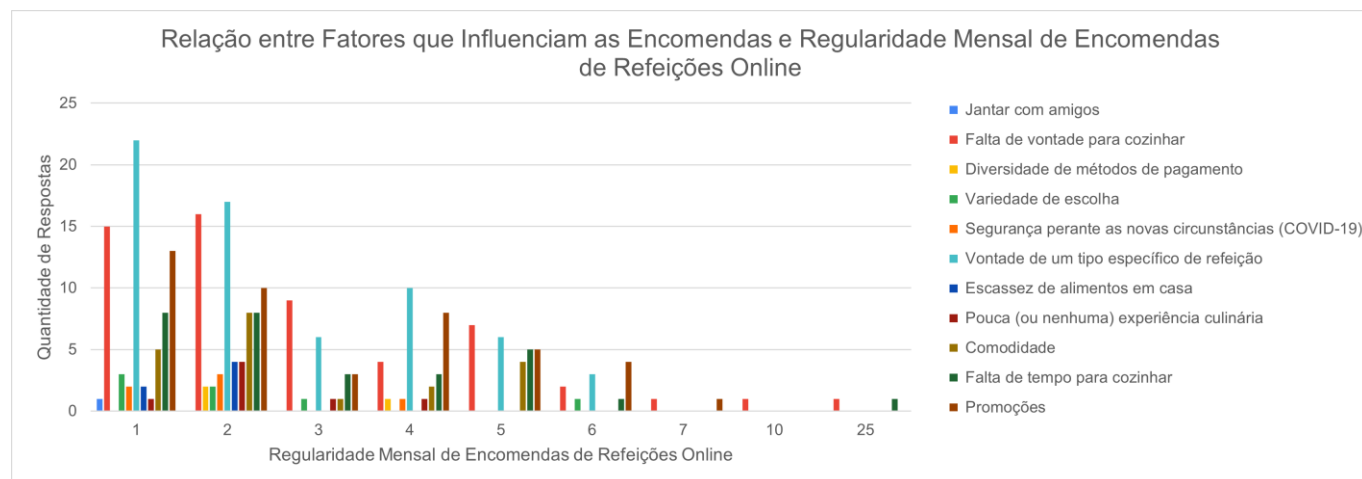


Figura 16: Relação entre Motivos que Influenciam as Encomendas e Regularidade de Encomendas de Refeições Online

A ilustração acima representa a Relação entre Motivos que Influenciam as Encomendas com a Regularidade de Encomendas de Refeições Online por parte da nossa amostra composta por 90 consumidores. Esta relação tem como objetivo o entendimento das razões que influênciam os consumidores a realizar encomendas de refeições online. Ao analisar os dados da amostra, podemos constatar que o motivo mais comum para quem encomenda entre 1 e 5 refeições online, mensalmente, é a “Vontade de um tipo específico de refeição”, outro motivo que é importante salientar é que independentemente da regularidade mensal de encomendas, a “Falta de vontade de cozinhar” é um motivo comum em todos os participantes. As “Promoções” também têm um peso significativo nos consumidores em questão.

2.3.9 Relação entre Entraves às Encomendas e Regularidade de Encomendas de Refeições Online

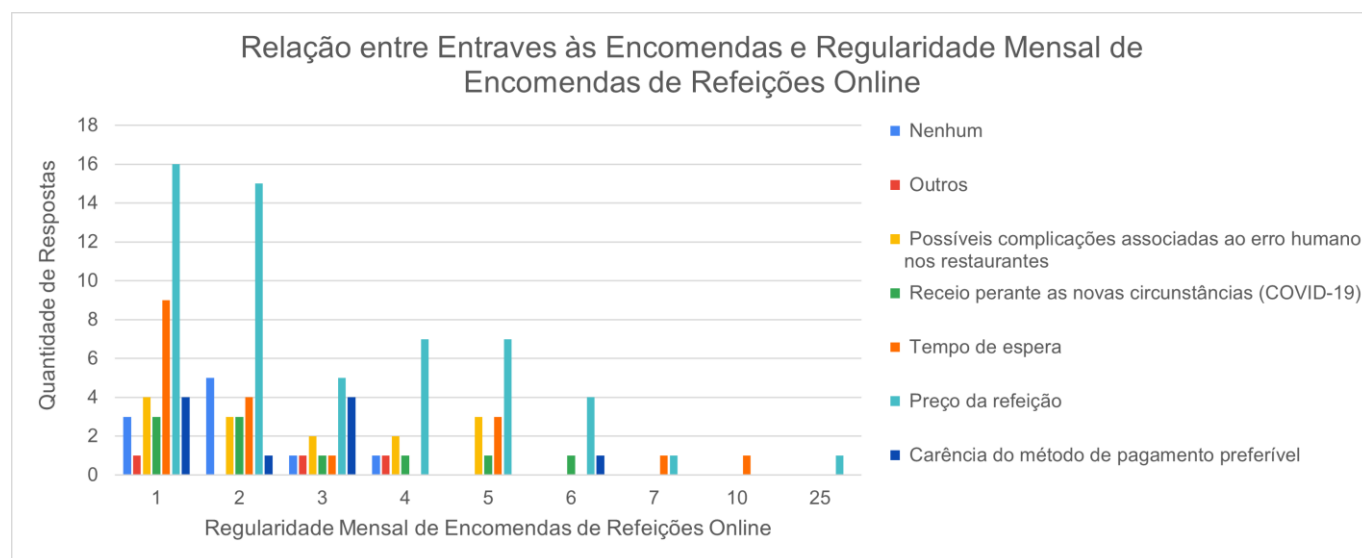


Figura 17: Relação entre Entraves às Encomendas e Regularidade de Encomendas de Refeições Online

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

Observando as 90 respostas consideradas no estudo, observamos que mais de metade da amostra (61.5% correspondentes a 56 inquiridos) consideram o preço da refeição o principal entrave à encomenda de uma refeição online, existindo, no entanto, vários outros motivos de entre os quais destacamos:

- Tempo de espera, entrave partilhado por 20 inquiridos;
- Possíveis complicações associadas ao erro humano nos restaurantes, partilhado por 14 inquiridos;
- Carência do método de pagamento preferível, partilhado por 11 inquiridos.

A ilustração acima compara os entraves às encomendas com a regularidade de encomendas de refeições online. Através da análise do gráfico, observamos que o entrave do preço da refeição se mantém constante, independentemente da regularidade, ao contrário dos restantes entraves que se encontram distribuídos aleatoriamente, não se apresentando de forma constante. Este fenómeno pode ser um indicador de que os valores que observamos no atributo “Regularidade de Encomendas Online” poderiam ser mais elevados, fosse o preço da refeição mais reduzido. No contexto da nossa amostra, podemos então observar que existe uma relação direta entre os entraves às encomendas e a regularidade mensal de encomendas de refeições online.

2.4 Verificação da qualidade dos dados

Com esta tarefa pretende-se examinar a qualidade dos dados, revendo-os com especial atenção a anomalias ou erros, colocando questões tais como: os dados estão completos (cobrem todos os casos necessários)? Os dados são corretos ou contêm erros e nesse caso, quão comum são eles? Existem valores em falta nos dados? Se sim, como se podem estes representar? Onde ocorrem e qual a sua importância?

Desta forma, para garantir a qualidade dos dados e evitar possíveis erros no futuro elaborámos um plano de solução:

- Os dados estão bem representados?
Sim, estão representados adequadamente e devidamente identificados.
- O significado dos atributos e os seus respetivos valores estão de acordo?
Sim, todos os atributos e respetivos valores estão em conformidade.
- Existem atributos em falta e campos em branco?
Não, não existem atributos em falta nem campos em branco.
- Existem atributos com diferentes valores que tenham significado idêntico?
Sim, esta situação ocorre.

Deparamo-nos com este problema na questão de possível resposta aberta, nomeadamente, na opção “Outra”, referente aos entraves - “Que entraves sente que se impõem quando pretende encomendar uma refeição online?” – onde foram obtidas duas respostas com significado idêntico:

1. “Valor das taxas de entrega”;
2. “Taxas de entregas elevadas”.

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

- Há escassez de valores?

Mediante a Ilustração x podemos verificar que em relação às idades existe grande tendência de valores, isto devido à grande aderência por parte das faixas etárias compreendidas entre os 15 e os 25 anos e à falta de adesão das restantes faixas etárias, provocando escassez de valores. Desta forma, a falta destes valores causa um grande impacto no estudo a ser realizado.

Sendo assim, não foi possível extrair observações acerca do fator Idade.

O mesmo acontece com o fator estado civil, onde existiu uma predominância de inquiridos solteiros e uma escassez de valores referentes a outros grupos, incapacitando-nos de obter conclusões sobre este fator dentro da nossa amostra.

- Há valores duplicados?
Qualquer valor é único.
- Os valores são plausíveis?
Todos os valores são plausíveis.

Pode-se garantir, depois de analisada toda a informação e verificando que todos os campos estão preenchidos de forma correta, que a informação é válida.

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

3. Preparação dos dados

A fase de preparação dos dados cobre todas as Atividades de construção do conjunto de dados final (dados que serão carregados na ferramenta de modelação) a partir do conjunto de dados usados recolhidos inicialmente. As tarefas de preparação dos dados poderão ser feitas por diversas vezes e não tem necessariamente uma ordem predefinida. As tarefas incluem a seleção das tabelas, dos registos e dos campos a usar, bem como a transformação e limpeza desses dados antes destes serem usados nas ferramentas de modelação.

3.1 Seleção dos dados

O processo de seleção de dados consiste na “filtragem” dos dados obtidos, através da compreensão do domínio e dos objetivos da tarefa, distinguindo assim os dados que serão sujeitos a análise, e os que serão ignorados.

O principal critério que utilizámos para a seleção de dados foi a influência dos dados no atributo “Regularidade de Encomendas Online”, no entanto, a qualidade dos mesmos revelou-se também um fator imprescindível a este processo.

De forma a ir de acordo com os referidos objetivos do estudo, escolhemos incluir na análise de estudo os seguintes dados que nos parecem ter mais influência no mercado online de entrega de refeições ao domicílio:

4. Idade
5. Género
6. Rendimento mensal do agregado familiar
7. Realiza encomendas de refeição online
8. Diferentes razões que influenciam a encomenda de refeições online:
 1. Vontade de um tipo específico de refeição
 2. Falta de tempo para cozinhar
 3. Falta de vontade para cozinhar
 4. Promoções
 5. Comodidade

3.2 Limpeza dos dados

Esta tarefa assume um papel importante no processo de preparação e futura análise dos dados pois é nesta fase que são excluídos todos dados considerados irrelevantes, desnecessários ou descontextualizados. Este processo assume um papel importantíssimo aquando da Verificação da Qualidade dos Dados.

Posto isto, neste processo de preparação de dados, vimos o nosso trabalho simplificado pelos nossos inquiridos, uma vez que a serenidade nas respostas, aliada à clareza das questões, quase não proporcionou resultados sujeitos a exclusão na análise.

A utilização do advérbio “quase”, deve-se à única resposta obtida que não se manteve nos limites das circunstâncias descritas e que, por essa mesma razão, decidimos não contabilizar na análise de dados: na questão de resposta aberta “Estado civil”, obtivemos a resposta “Relação Aberra”. Dentro das questões de possível resposta aberta ainda uma outra intervenção foi necessária, nomeadamente, na questão referente aos entraves onde foram obtidas duas respostas com significado idêntico. Decidimos, então, agrupar estas respostas num único atributo/tópico de nome “Outras”.

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

Quanto às decisões de exclusão ou ignoração deliberada tomadas por nós, facilmente explicamos a ausência dos restantes dados na análise através da diversidade influenciada pelas circunstâncias. Isto é, das 113 respostas obtidas, a grande maioria dos inquiridos insere-se nas mesmas circunstâncias académicas, etárias e sociais, invalidando a utilização desses dados (que poder-se-iam verificar pertinentes) por falta de diversidade de respostas, o que invalidaria a veracidade do estudo feito, por se restringir a um coletivo pouco abrangente.

Escolhemos ainda não considerar a resposta “0”, dada à pergunta de resposta aberta “Com que regularidade encomenda mensalmente refeições online?”, por entendimento coletivo de que nada acrescentaria ao estudo em questão.

3.3 Construção dos dados

Alguns algoritmos trabalham apenas com valores numéricos e outros apenas com valores categóricos. Nestes casos, é necessário transformar os valores numéricos em categóricos ou os categóricos em valores numéricos. Não existe um critério único para transformação dos dados e diversas técnicas podem ser usadas de acordo com os objetivos pretendidos. Algumas das técnicas empregadas nesta etapa são: agrupamento (agrupa valores em faixas sumarizadas), generalização (converte valores muito específicos para valores mais genéricos), normalização (colocar as variáveis em uma mesma escala) e a criação de novos atributos, designados por atributos derivados (gerados a partir de outros já existentes).

Nesta etapa do projeto, recorreremos à técnica de agrupamento, onde agrupámos as faixas etárias em 5 classes distinguidos pelas letras A (dos 15 aos 20 anos), B (dos 21 aos 25 anos), C (dos 26 aos 30 anos), e a letra D para todas as idades acima dos 30 anos. Como não havia respostas para a faixa etária abaixo dos 15anos, e como não fazia sentido para o tema em questão (pois pessoas inseridas nessa faixa etária ainda não possuem independência para a realização de encomendas de refeições online), decidiu-se não implementar a respetiva classe.

Idade		
Limite inferior	Limite superior	Classe
15	20	A
21	25	B
26	30	C
30		D

Tabela 3: Categorias da Idade

Para além desta situação, voltámos ainda a recorrer a este método para agrupar os vários rendimentos mensais do agregado familiar dos inquiridos. Para tal, agrupámos desta vez em 6 classes, novamente distinguidos pelas letras A (para rendimentos inferiores a 600€ mensais), B (para rendimentos mensais entre os 600€ e os 800€), C (para rendimentos mensais entre os 800€ e os 1200€), D (para rendimentos mensais entre os 1200€ e os 1600€), E (para rendimentos mensais entre os 1600€ e os 2000€), e a letra F (para rendimentos superiores a 2000€ mensais).

Rendimento mensal do agregado familiar		
Limite inferior	Limite superior	Classe
	600 €	A
600 €	800 €	B
800 €	1200 €	C
1200 €	1600 €	D

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

1600 €	2000 €	E
2000 €		F

Tabela 4: Categorias do Rendimento mensal do agregado familiar

Para além dos dois já referidos agrupamentos, recorreremos ainda à criação de novos atributos derivados, isto é, a criação de novos campos construídos através de um outro já existente. Este processo derivou da necessidade de individualizar os principais fatores que levam os inquiridos a encomendar uma refeição, destacando os 5 fatores mais escolhidos pelo universo da amostra. Para isto, foram criados atributos individuais, derivados do atributo “Quais os principais fatores que o levam a encomendar uma refeição?”. Assim, elaborámos 5 novos atributos, “vontade de um tipo específico de refeição”, “Falta de tempo para cozinhar”, “Falta de vontade para cozinhar”, “Promoções” e “Comodidade”, todos com o intuito de clarificar se o inquirido selecionou, ou não, a opção em questão, e, para tal, todos com o campo único de “sim” ou “não”.

3.4 Integração dos dados

Como já referido, os dados provêm única e inteiramente de um inquérito elaborado na sua integra pelos constituintes do grupo, e respondido por um público generalizado não específico. Sendo assim, não existe a necessidade da integração destes dados pois já possuímos um repositório único e consistente.

Uma vez preenchidos os inquéritos por uma amostra considerável, os dados foram exportados para a ferramenta Microsoft Excel de modo a trabalhar neles com uma abordagem simplificada e mais intuitiva.

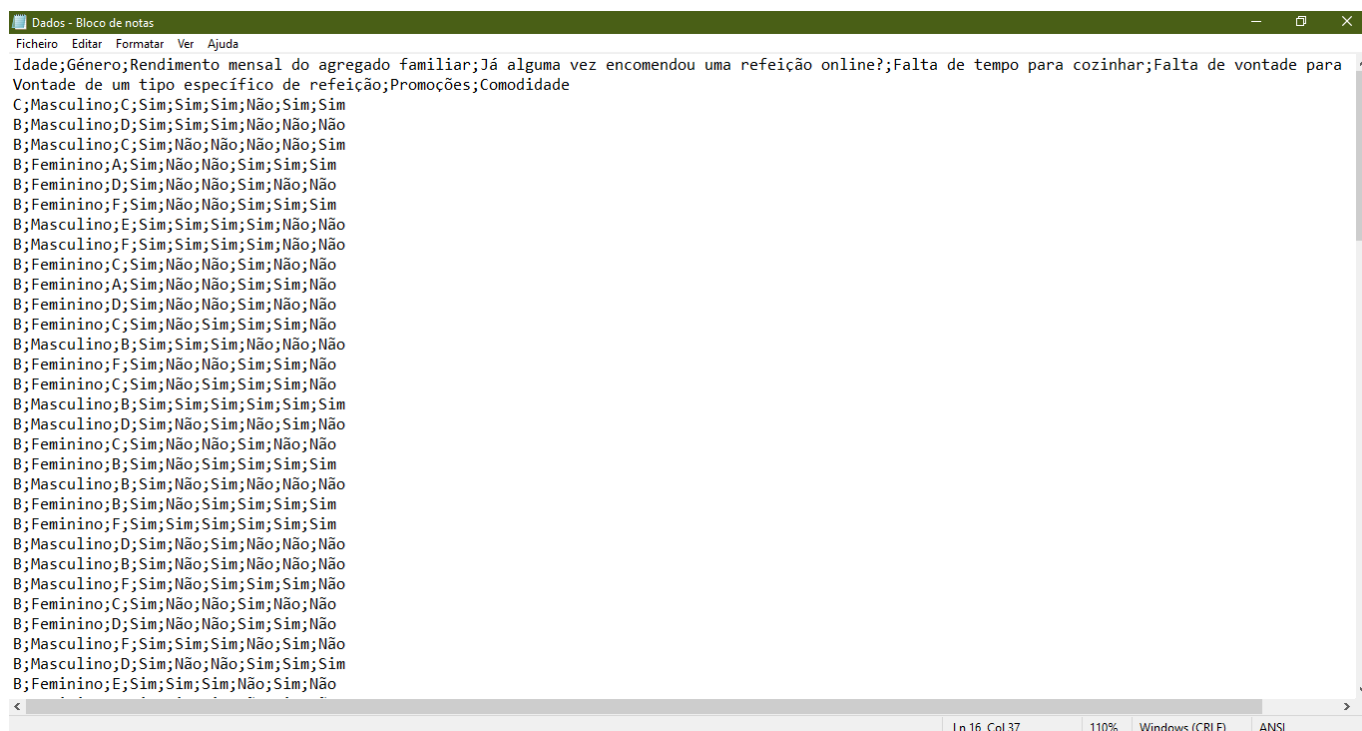
3.5 Formatação dos dados

Esta tarefa consiste na alteração sintática feita nos dados, necessária para a ferramenta de modelação a usar, mas sem alterar o seu significado. Necessitámos, então, de recorrer a alterações sintáticas em três ocasiões.

As primeiras duas situações foram bastante semelhantes e ambas se deveram à especificidade do algoritmo de associação “à priori”. A presença de caracteres numéricos não permitiu a interpretação dos dados da parte do algoritmo, como tal, e também aliada à justificação já descrita no ponto 3.3 - Construção de dados, as classes de faixas etárias, bem como as classes de rendimento mensal do agregado familiar, foram organizadas em classes crescentes, denominadas com as primeiras letras do alfabeto. Assim, com a atribuição de letras às classes apresentadas, o algoritmo de associação “à priori” foi finalmente capaz de interpretar os dados de ambos estes campos e, desta forma, proceder à análise dos mesmos.

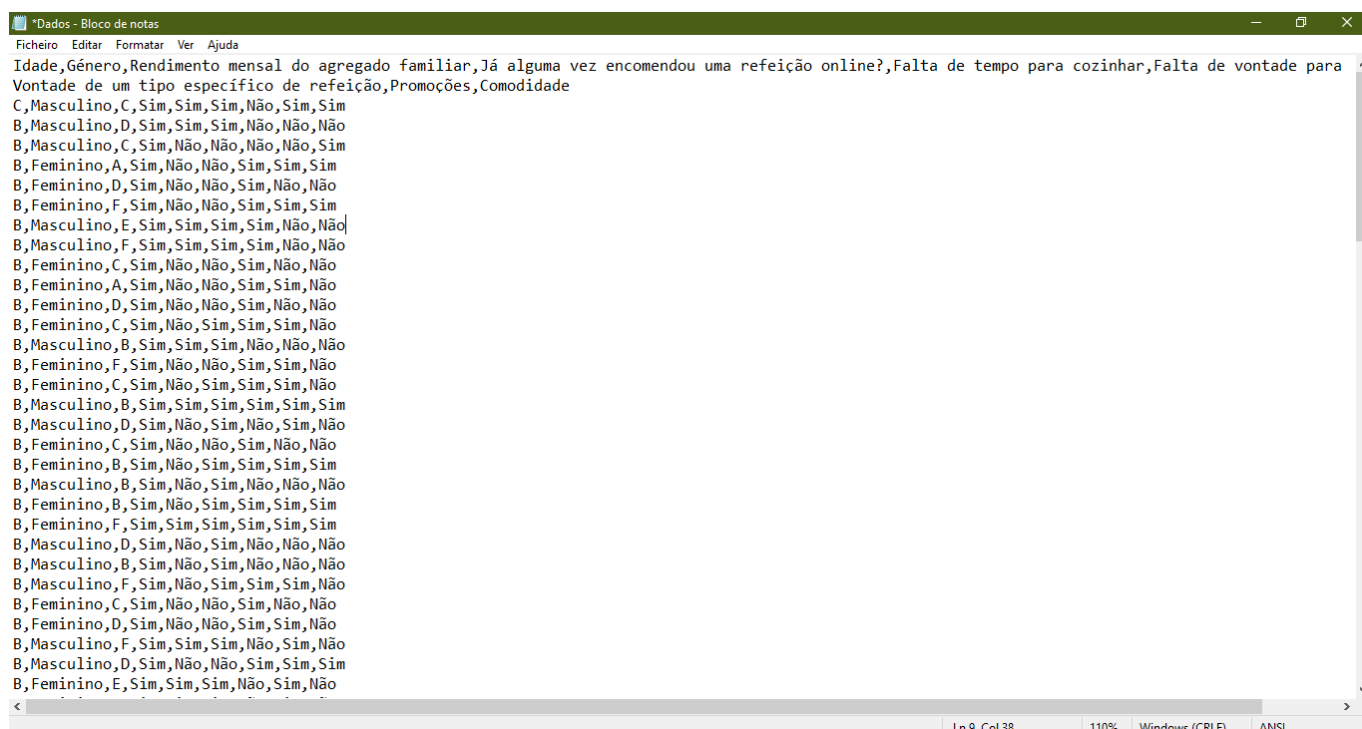
A terceira situação adveio das características do programa de mineração de dados Weka Explorer, que apresentava a vírgula como separador de campos por definição, ao contrário do programa utilizado para trabalhar os dados adquiridos, Microsoft Excel, que tinha por definição o ponto-e-vírgula como separador de campos. Para isto, servimo-nos do programa Bloco de Notas, do Windows OS, e através da ferramenta substituir, alterámos todos os ponto-e-vírgulas gerados automaticamente pelo Microsoft Excel e substituímos pela vírgula necessária ao programa Weka Explorer.

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio



A screenshot of a text editor window titled "Dados - Bloco de notas". The editor displays a CSV file with unformatted data. The first row contains headers: "Idade;Género;Rendimento mensal do agregado familiar;Já alguma vez encomendou uma refeição online?;Falta de tempo para cozinhar;Falta de vontade para Vontade de um tipo específico de refeição;Promoções;Comodidade". Subsequent rows contain data points separated by semicolons, such as "C;Masculino;C;Sim;Sim;Sim;Não;Sim;Sim" and "B;Masculino;D;Sim;Sim;Sim;Não;Não;Não". The status bar at the bottom indicates "Ln 16, Col 37", "110%", "Windows (CRLF)", and "ANSI".

Figura 18: Ficheiro CSV antes da formatação necessária



A screenshot of a text editor window titled "*Dados - Bloco de notas". The editor displays the same CSV data as Figure 18, but now each data row is enclosed in double quotes, and the semicolons are escaped with backslashes. For example, the first data row is now "\"C;Masculino;C;Sim;Sim;Sim;Não;Sim;Sim\"". The status bar at the bottom indicates "Ln 9, Col 38", "110%", "Windows (CRLF)", and "ANSI".

Figura 19: Ficheiro CSV depois da formatação necessária

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

4. Modelação

Nesta fase, a fase de modelação, são seleccionadas e aplicadas as técnicas de Data Mining mais apropriadas e os seus parâmetros são ajustados de forma a otimizar os resultados. Normalmente, existem várias técnicas para o mesmo tipo de problema de DM, sendo que algumas têm pressupostos específicos sobre a forma como os dados são apresentados (qualidade, formatação, distribuição), por isso, pode ser necessário voltar à fase de preparação de dados.

4.1 Seleção da técnica de modelação

No âmbito deste trabalho, optou-se por adotar as seguintes técnicas: Segmentação e Associação, usando o software WEKA para implementação destas.

4.1.1 Segmentação

A segmentação, mais conhecida por agrupamento, é um processo de partição de uma população heterogénea em vários subgrupos ou segmentos (clusters) mais homogéneos, isto é, que partilham tendências e padrões semelhantes. Em geral, as medidas de similaridade usadas são as de distâncias tradicionais sendo esta a medida utilizada no algoritmo escolhido para a realização da tarefa de agrupamento. Os elementos de um cluster são considerados similares aos elementos no mesmo cluster (similaridade intra-grupo) e dissimilares aos elementos nos outros clusters (similaridade inter-grupos). Na segmentação, não há classes predefinidas, os registos são agrupados de acordo com a semelhança, o que a diferencia da tarefa de classificação.

Algoritmo SimpleKMeans - Usa o conceito da centroide (centro geométrico). Dado um conjunto de dados, o algoritmo seleciona de forma aleatória k registos, cada um representando um agrupamento. Para cada registo restante, é calculada a similaridade entre o registo analisado e o centro de cada agrupamento. O objeto é inserido no agrupamento com a menor distância, ou seja, maior similaridade. O centro do cluster é recalculado a cada novo elemento inserido.

4.1.2 Associação

A tarefa de associação consiste em determinar o relacionamento dos atributos mais frequentes em determinado conjunto de dados. Apresenta a forma: SE atributo X ENTÃO atributo Y. Esta construção recebe o nome de Regra de Associação (Association Rules).

O exemplo clássico é a determinação dos produtos que costumam ser colocados juntos num carrinho de supermercado, daí o termo 'análise de market basket'. As cadeias de distribuição utilizam a associação para planejar a disposição dos produtos nas prateleiras das lojas ou num catálogo, de modo que os itens geralmente adquiridos na mesma compra sejam vistos próximos entre si (exemplo: SE compra leite e pão TAMBÉM compra manteiga).

Algoritmo Apriori – Vai de encontro à estratégia de mineração de itens frequentes (Frequent Itemset Mining) que pode ser visualizada em duas etapas: primeiro, um conjunto de itens frequentes (Frequent Itemset) é criado, respeitando um valor mínimo de frequência para os itens. Depois, as regras de associação são geradas pela mineração desse conjunto. Para garantir resultados válidos, os conceitos de suporte e confiança são utilizados em cada regra produzida.

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

A medida de suporte indica o percentual de registos (dentre todo o conjunto de dados) que se encaixam nessa regra. Já a confiança mede o percentual de registos que atendem especificamente a regra, por exemplo, o percentual de quem compra leite e pão e também compra manteiga. Para uma regra ser considerada forte, ela deve atender a um certo grau mínimo de suporte e confiança.

4.2 Geração do desenho de testes

Antes da construção do modelo, deve ser criado um procedimento ou mecanismo para testar a qualidade do modelo e a sua validade. Normalmente, separa-se o conjunto de dados em dois subconjuntos, um de treino do modelo e outro para ser usado como teste desse modelo onde este último é ainda dividido num subconjunto de teste de validação. Os dados de treino serão usados para construir o modelo, enquanto que os dados de teste serão usados posteriormente para estimar a qualidade do mesmo.

Essa divisão em grupos é necessária para que o modelo não fique dependente de um conjunto de dados específico e, ao ser submetido a outros conjuntos (com valores diferentes dos usados na construção do modelo), não apresente resultados insatisfatórios. Pretendemos, assim, evitar a possibilidade de perder precisão para outros conjuntos de dados à medida que se aumenta a precisão do modelo para um conjunto de dados específico (dados de treino).

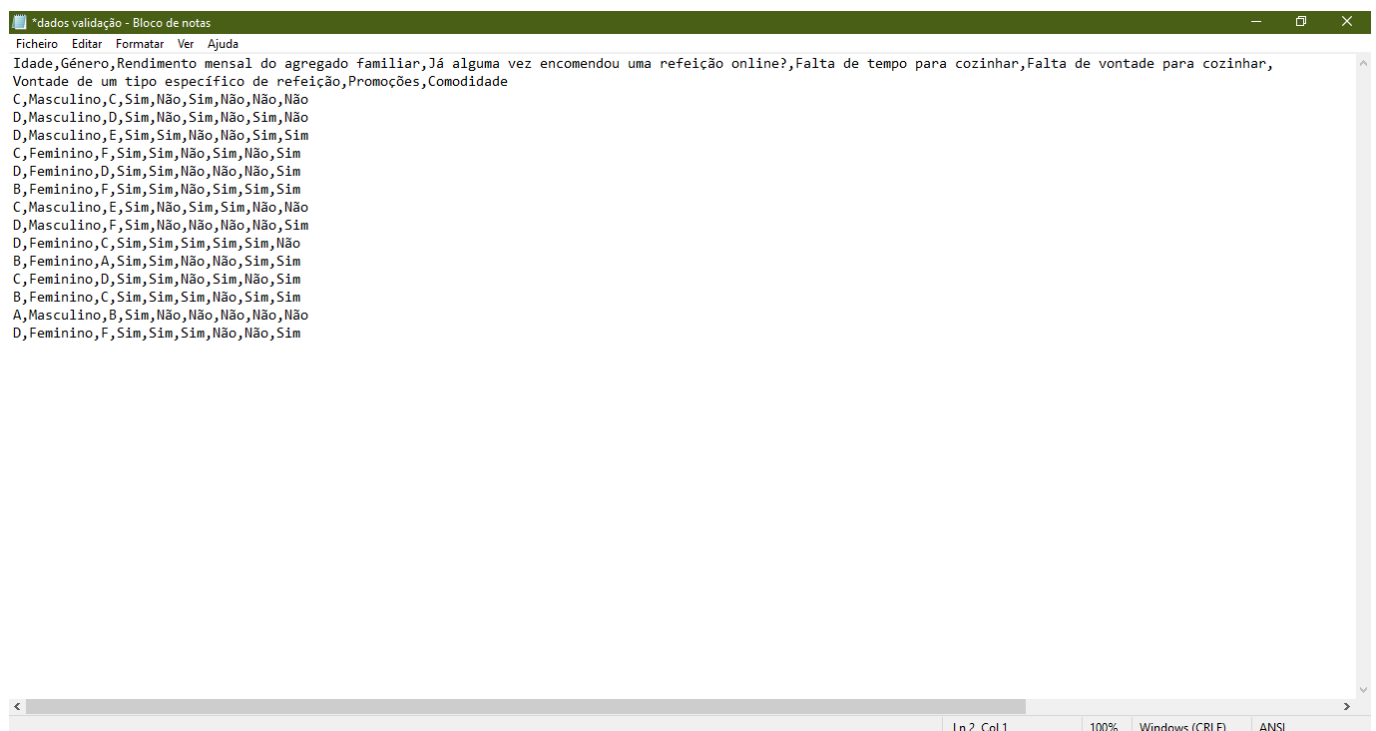


Figura 20: Ficheiro CSV para dados de validação

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

4.3 Construção do modelo

O objetivo agora é usar a ferramenta de modelação sobre o conjunto de dados preparados de forma a construir os modelos de Agrupamento e Associação e, consequentemente, interpretar os vários resultados obtidos.

Em resultado extraído do algoritmo “SIMPLEKMEANS” obtivemos um modelo de agrupamento com 2 perfis, sendo, assim, possível entender qual o consumidor típico e ainda quais os 2 principais segmentos. Importa referir que a tarefa de agrupamento separa os registos em subconjuntos de tal forma que os registos dum agrupamento partilhem propriedades comuns, distinguindo-os de outros agrupamentos, maximizando a similaridade intra-cluster e minimizando a similaridade inter-cluster.

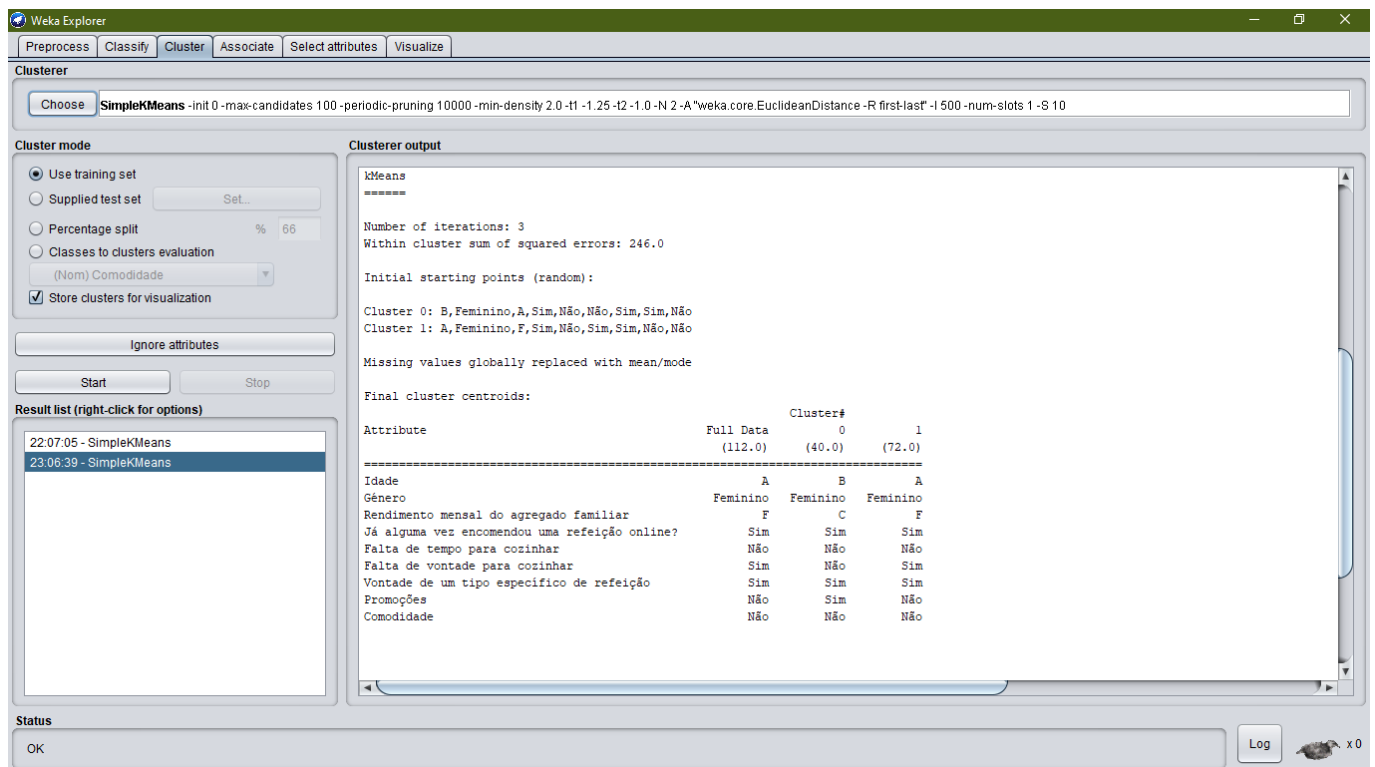


Figura 21: Output do algoritmo SIMPLEKMEANS

O consumidor típico de refeições encomendadas online é definido da seguinte forma: idade entre os 15 e 20 anos, do género feminino, com rendimento mensal do agregado familiar superior a 2000€, não possui falta de tempo para cozinhar, mas, por outro lado, detém de falta de vontade para cozinhar, susta de vontade de um tipo específico de refeição, não sendo influenciado por promoções nem pela comodidade oferecida pelos serviços de entrega de refeições online.

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

Perfil do consumidor Típico	
Idade	Entre 15 e 20 anos (A)
Género	Feminino
Rendimento mensal do agregado familiar	>2000€ (F)
Falta de tempo para cozinhar	Não
Falta de vontade para cozinhar	Sim
Vontade de um tipo específico de refeição	Sim
Promoções	Não
Comodidade	Não

Tabela 5: Perfil do consumidor Típico

O primeiro segmento/cluster, representado como cluster 0, caracteriza 36% dos consumidores da seguinte forma: idade entre os 21 e 25 anos; género feminino; rendimento mensal do agregado familiar entre 800€ e 1200€; não possui falta de tempo para cozinhar, nem falta de vontade para tal; detém vontade de um tipo específico de refeição; leva em conta as promoções do método de encomenda online aquando da decisão da refeição; não é persuadido pela comodidade oferecida pelos serviços de entrega de refeições online.

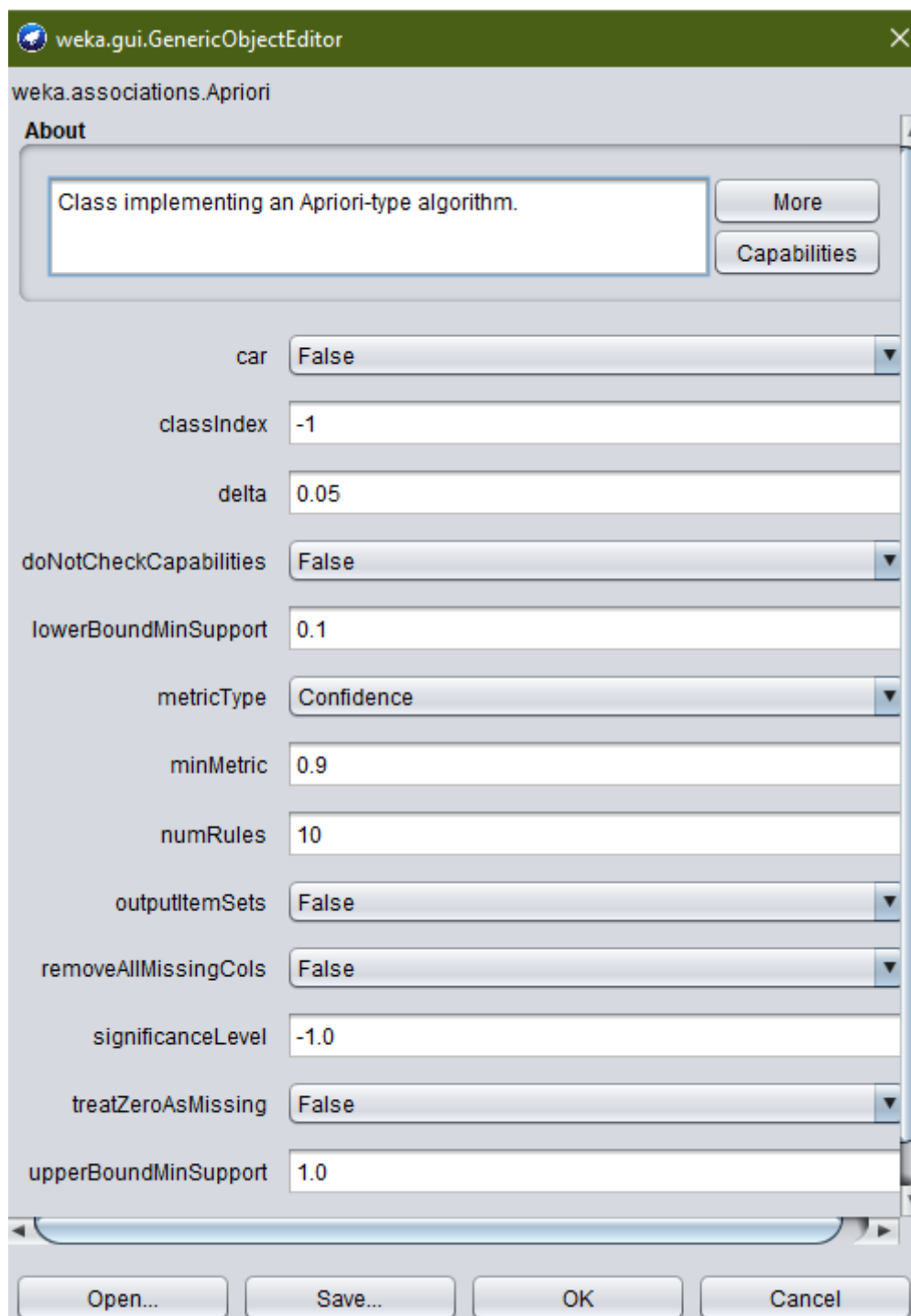
Cluster 0	
Idade	Entre 21 e 25 anos (B)
Género	Feminino
Rendimento mensal do agregado familiar	Entre 800€ e 1200€ (C)
Falta de tempo para cozinhar	Não
Falta de vontade para cozinhar	Não
Vontade de um tipo específico de refeição	Sim
Promoções	Sim
Comodidade	Não

Tabela 6: Perfil obtido pelo cluster 0

O segundo segmento/cluster, representado como cluster 1, caracteriza 64% dos consumidores de acordo com o consumidor típico.

Analizados os resultados obtidos pelo modelo de Agrupamento, iremos então efetuar a devida análise à técnica de Associação e perceber quais as conclusões a retirar da mesma, logo após a listagem dos parâmetros utilizados.

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio



The image shows a screenshot of the Weka GUI's GenericObjectEditor window for the Apriori algorithm. The window title is 'weka.gui.GenericObjectEditor'. The class name is 'weka.associations.Apriori'. The 'About' tab is selected, showing a description: 'Class implementing an Apriori-type algorithm.' with 'More' and 'Capabilities' buttons. Below this, various parameters are listed with their current values:

Parameter	Value
car	False
classIndex	-1
delta	0.05
doNotCheckCapabilities	False
lowerBoundMinSupport	0.1
metricType	Confidence
minMetric	0.9
numRules	10
outputItemSets	False
removeAllMissingCols	False
significanceLevel	-1.0
treatZeroAsMissing	False
upperBoundMinSupport	1.0

At the bottom of the window are four buttons: 'Open...', 'Save...', 'OK', and 'Cancel'.

Figura 22: Parâmetros utilizados

Um dos mais importantes parâmetros é o da confiança que neste caso está em 0.9, ou seja, em 90%. A confiança numa regra $X \rightarrow Y$ significa a proporção de instâncias em que se verifica a condição X que também são do tipo Y.

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

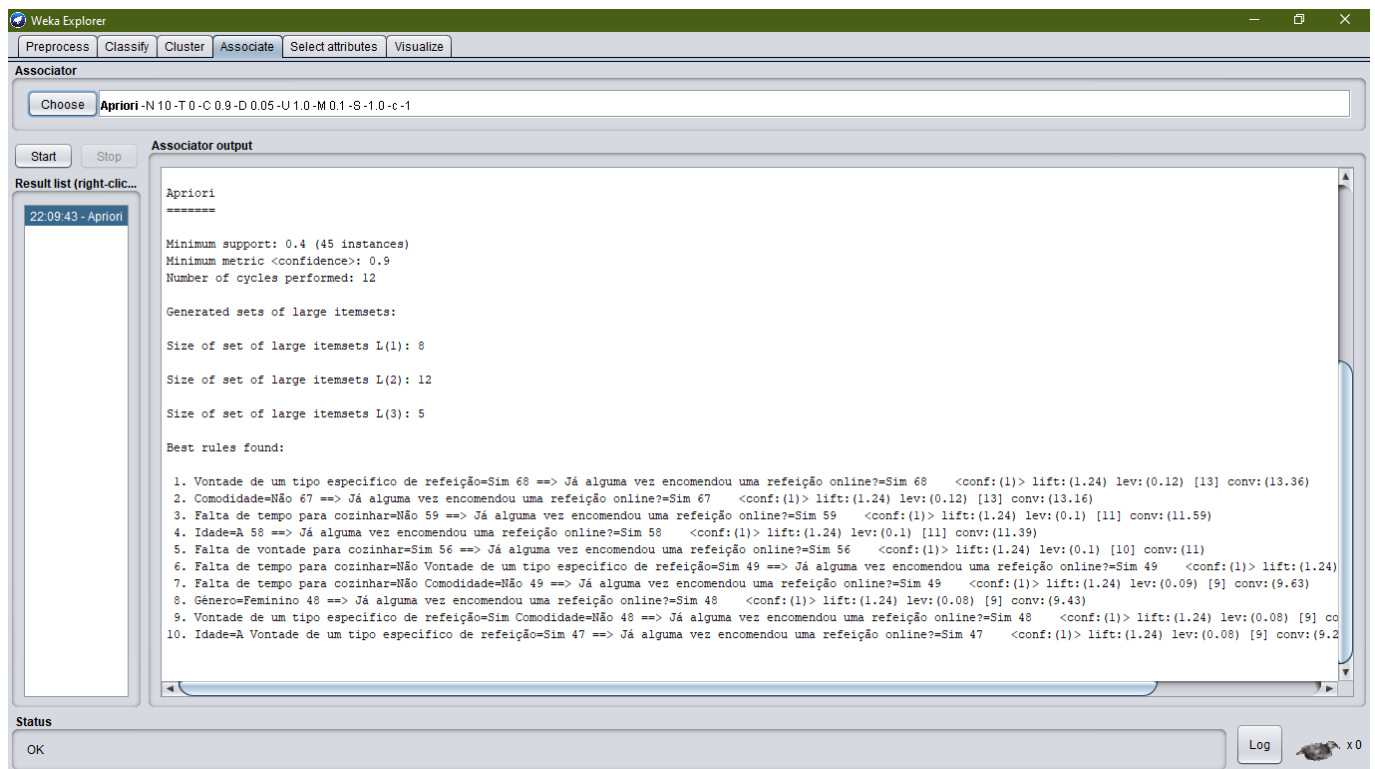


Figura 23: Output do modelo APRIORI

Após execução do algoritmo “APRIORI”, foi feito o pós-processamento adequado aos resultados do modelo:

- 68 pessoas de 90 que responderam ao inquérito encomendaram uma refeição online porque tinham vontade de um tipo específico de refeição.
Suporte de 68/90, ou seja, de 76%;
- 67 pessoas de 90 que responderam ao inquérito não encomendaram uma refeição online devido há comodidade.
Suporte de 67/90, ou seja, de 74%;
- 59 pessoas de 90 que responderam ao inquérito não encomendaram uma refeição online devido há falta de tempo para cozinhar.
Suporte de 59/90, ou seja, de 66%;
- 58 pessoas de 90 que responderam ao inquérito encomendaram uma refeição online têm uma idade compreendida entre 15 e 20 anos.
Suporte de 58/90, ou seja, de 64%;
- 56 pessoas de 90 que responderam ao inquérito encomendaram uma refeição online devido há falta de vontade de cozinhar.
Suporte de 56/90, ou seja, de 62%;

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

- 49 pessoas de 90 que responderam ao inquérito encomendaram uma refeição online não tendo falta de tempo para cozinhar, mas por ter vontade de um tipo específico de refeição. Suporte de 49/90, ou seja, de 54%;
- 49 pessoas de 90 que responderam ao inquérito encomendaram uma refeição online não o fazendo pela falta de tempo para cozinhar nem por comodidade. Suporte de 49/90, ou seja, de 54%;
- 48 pessoas de 90 que responderam ao inquérito encomendaram uma refeição online são do sexo feminino. Suporte de 48/90, ou seja, de 53%;
- 48 pessoas de 90 que responderam ao inquérito encomendaram uma refeição online porque tinham vontade de um tipo de refeição e não pela comodidade. Suporte de 48/90, ou seja, de 53%;
- 47 pessoas de 90 que responderam ao inquérito encomendaram uma refeição online porque tinham vontade de um tipo de refeição e tinham uma idade compreendida entre 15 e 20 anos. Suporte de 47/90, ou seja, de 52%;

Foram dispostas acima as 10 regras mais significativas pelo modelo, onde todas elas detêm uma confiança de 100%.

Analisando os resultados do modelo, observa-se que existem algumas regras que fazem mais sentido à luz do conhecimento que se tem do tema.

Destaca-se, assim, a vontade por um tipo de refeição específico (76%), a idade do consumidor compreendida entre os 15 e 20 anos (64%), a falta de vontade para cozinhar (62%) e o género feminino (53%).

Uma vez identificados os grupos no modelo de agrupamento, identificadas as regras no modelo de associação e, consequentemente, feita a interpretação dos mesmos, falta agora dar resposta à principal questão em estudo: quais os principais fatores que influenciam a opção pela encomenda de refeições online?

Efetuando uma avaliação aos resultados dos modelos, constatámos que o fator com maior destaque foi a vontade por um tipo específico de refeição, seguido pelo fator que dita a falta de vontade para cozinhar. Ainda é possível destacar as promoções apresentadas pelo método de encomenda online como um fator influente, embora em menor escala (36% dos casos), na opção pela encomenda de refeições online.

Por último, realçar a predominância do género feminino, concluindo, assim, que são as mulheres com a maior tendência a optar pela encomenda de refeições online.

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

4.4 Revisão do modelo

O modelo deverá agora ser revisto para que se garanta que cumpre os critérios de avaliação de sucesso anteriormente estabelecidos para a mineração de dados. O modelo deverá agora ser revisto para que se garanta que cumpre os critérios de avaliação de sucesso anteriormente estabelecidos para a mineração de dados.

Para que tal possa ser efetuado, iremos analisar os resultados obtidos, retirando as conclusões devidas relativamente às qualidades do modelo e procedendo à hierarquização das mesmas.

Assim, após uma análise refletiva do modelo obtido, observamos que, em concordância com o universo amostral do nosso inquérito, o fator que mais influencia a encomenda de uma refeição online é a vontade de um tipo específico de refeição.

Dado o sucesso e qualidade do modelo desenvolvido, conseguimos ainda observar os padrões do consumidor mais comum do mercado de restauração online, possuindo informações concretas relativas à sua idade e sexo, campos considerados mais relevantes no universo do estudo.

Como a nossa amostra de consumidores com faixa etária diferente à predominante é pouco significativa, devido à falta aderência por parte desta, não podemos concluir se a constatada predominância das idades influencia a regularidade de encomendas de refeições online.

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

5. Avaliação

Nesta fase do projeto já foi construído o modelo (ou modelos) que aparentemente tem qualidade numa perspetiva da análise dos dados. É importante fazer uma avaliação do modelo e rever os passos executados para construir o modelo, para que haja uma certeza de que este modelo está de acordo com os critérios de sucesso do negócio. Um dos aspetos mais importantes passa por ter a certeza de que não existe nenhum aspeto do negócio que não tenha sido devidamente considerado. No fim desta fase, dever-se-á decidir do uso dos resultados de DM.

5.1 Avaliação de resultados

A fase de avaliação contempla em primeiro lugar fatores tais como precisão e generalidade do modelo. Este passo avalia o grau de cumprimento dos Objetivos do negócio por parte do modelo e procura descobrir se existem alguns razões de negócio que justifiquem alguma deficiência do modelo. Outra área de avaliação consistiria no teste do modelo no ambiente real, com uma aplicação real, caso o tempo e o orçamento o permitam, avaliação essa que não iremos implementar como já tinha sido mencionado anteriormente.

Por outro lado, esta fase pretende cobrir outros aspetos, tal como a avaliação dos resultados gerados com a mineração dos dados. Os resultados da mineração dos dados usam modelos que estão necessariamente relacionados com os Objetivos de negócio originais e outros resultados que poderão não estar relacionados com esses Objetivos originais, mas que poderão representar uma oportunidade de mudança de direção e dos Objetivos de negócio.

Após interpretação dos resultados podemos concluir que cumprimos os objetivos do Data Mining, pois foi possível perceber, através desses mesmos resultados dos modelos, o motivo que leva aos consumidores optar pela entrega de refeições ao domicílio. Relativamente ao critério de sucesso do tema/negócio, ou seja, caso as receitas provenientes das vendas efetuadas pela encomenda online efetivamente aumentarem, ou, na pior das hipóteses, se mantenha o número e não haja reduções, apenas é possível verificar em que medida este critério é atingido caso houvesse implantação dos resultados por parte dos restaurantes, sendo esta avaliação feita pela direção dos respetivos restaurantes.

Os nossos resultados iriam então ditar que medidas adotar de modo a alcançar o critério de sucesso. A implantação poderia ser feita através de campanhas publicitárias seguindo um plano baseado nos resultados do nosso estudo. Dessa forma conseguiríamos obter provas empíricas e perceber o impacto causado, confirmando, assim, se os resultados que obtivemos através da mineração de dados são ou não viáveis no mercado de restauração e fazer a respetiva aprovação dos modelos.

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

5.2 Revisão do processo

Nesta fase, já com dois modelos que satisfazem globalmente as necessidades do negócio, será então adequado fazer uma revisão mais detalhada do processo global de mineração feito até aqui no sentido de determinar se existe algum fator ou tarefa importante que tenha sido menosprezada. Esta revisão deverá também preocupar-se com aspetos de garantia da qualidade, como por exemplo, será que se construiu corretamente o modelo? Usaram-se apenas os campos para os quais havia autorização? Esses campos estarão disponíveis no futuro?

O processo de mineração de dados foi bastante desafiador, desde o entendimento do negócio à avaliação de resultados, mas com especial destaque na fase de preparação de dados onde vários entraves surgiram.

Primeiramente, na separação dos dados obtidos através dos inquéritos, onde, por vezes, apareciam unificados na mesma classe devido à hipótese de múltipla seleção da pergunta em questão. De modo a solucionar este problema, optámos por separar cada opção e contar o número de vezes em que esta foi selecionada pelos inquiridos.

Uma possível ação alternativa de modo a evitar este obstáculo passaria por na construção do inquérito, recorrer a perguntas isoladas de forma a que os dados se encontrassem previamente separados na hora da exportação para o Excel.

Outro entrave com que nos deparámos foi a formatação do ficheiro CSV requerido pelo weka, onde as classes deveriam se apresentar separadas por vírgulas, o que ao fazer a conversão a partir do ficheiro Excel (onde originariamente trabalhámos os dados) não acontecia. Solução para este problema já referida em tópicos anteriores.

Referente aos modelos, na nossa ótica estes foram bem construídos e os seus resultados, implementados de forma adequada, podem ajudar a atingir os critérios de sucesso do tema em questão.

5.3 Determinar os próximos passos

De acordo com os resultados da avaliação e a revisão do processo, damos como concluído o projeto.

Desta forma, decidimos concluir o estudo sem avançar com uma tentativa de intervenção ou participação nas tomadas de decisão dos inúmeros fornecedores diretos deste mercado de concorrência perfeita.

Numa fase posterior de implementação, no contexto de um mercado local e de pequena dimensão, como objetivo do crescimento e melhoria dos seus constituintes, o primeiro passo passaria pela divulgação do estudo aos interessados, para que o pudessem analisar e ter em consideração na próxima abordagem ao mercado.

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

6. Conclusão

Este projeto, não só nos proporcionou uma nova abordagem aos sistemas baseados em conhecimento, como uma perspectiva diferente da importância da análise de dados em qualquer projeto. Desta forma, e tendo em conta a complexidade e o detalhe que o trabalho em questão exige, sentimos que apesar de todas as adversidades, a pesquisa realizada, aliada aos conhecimentos adquiridos em contexto de aula, se revelaram cruciais no cumprimento dos objetivos delineados.

Resumindo o projeto, conseguimos reunir um número considerável de dados, recolhidos a partir de um inquérito com o intuito de entender quais os fatores que levam os consumidores a encomendar refeições online com entrega ao domicílio, de modo a auxiliar as empresas de restauração a lidar com a nova realidade, caso optassem por esta solução. Para tal construímos dois modelos, de agrupamento e de associação. No fim de toda a análise podemos concluir que, no contexto do universo da amostra em questão, em ambos os modelos obtivemos conclusões idênticas em relação aos fatores que se revelaram mais pertinentes nos consumidores do mercado online de entrega de refeições ao domicílio, sendo eles (por ordem de influência):

1. Vontade de um tipo específico de refeição
2. Falta de vontade para cozinhar
3. Promoções

Realçar ainda a predominância do género feminino, concluindo, assim, que são as mulheres com a maior tendência a optar pela encomenda de refeições online.

A informação obtida através do estudo possibilita uma oportunidade de melhorar o mercado online de entrega de refeições online, podendo ajudar algumas empresas do setor de restauração que tencionem embarcar neste mercado. No entanto, não conseguimos prever o impacto real que teria, como já explicado previamente.

Chegada a data de entrega e consequente conclusão do trabalho, gozamos da satisfação de poder fazer uma retrospectiva positiva do trabalho desenvolvido e levar connosco uma série de conhecimentos enriquecedores que iremos certamente utilizar, não só em futuros projetos educativos, como também no mercado de trabalho.

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

Anexos

Anexo 1 – Inquérito

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

O presente projeto aborda a análise dos diversos fatores que influenciam o consumo através do mercado de restauração online.

***Obrigatório**

Idade *

- ☐ < 15
- ☐ 15-20
- ☐ 21-25
- ☐ 26-30
- ☐ > 30

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

Género *

☐ Masculino

☐ Feminino

☐ Outro: _____

Rendimento mensal do agregado familiar *

☐ < 600€

☐ 600€ - 800€

☐ 800€ - 1200€

☐ 1200€ - 1600€

☐ 1600€ - 2000€

☐ > 2000€

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

Nível de Escolaridade *

- ☐ 1º Ciclo
- ☐ 2º Ciclo
- ☐ 3º Ciclo
- ☐ Ensino Secundário
- ☐ Licenciatura
- ☐ Mestrado
- ☐ Doutoramento
- ☐ Outro: _____

Ocupação *

- ☐ Estudante
- ☐ Trabalhador
- ☐ Estatuto de trabalhador estudante
- ☐ Desempregado

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

Estado Civil *

☐ Solteiro

☐ Casado

☐ Divorciado

☐ Viúvo

☐ Outro: _____

Já alguma vez encomendou uma refeição online? *

☐ Sim

☐ Não

Próxima

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

O presente projeto aborda a análise dos diversos fatores que influenciam o consumo através do mercado de restauração online.

Com que regularidade encomenda mensalmente refeições online? *

Sua resposta

Que métodos costuma utilizar para encomendar? *

Selecione todas as opções com que se identifique

- ☐ Aplicações tipo uber eats, glovo, etc
- ☐ Aplicação da cadeia de restauração
- ☐ Contactar o restaurante através do contacto telefónico

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

Quais os principais fatores que o levam a encomendar uma refeição? *

Selecione todas as opções com que se identifique

- ☐ Falta de tempo para cozinhar
- ☐ Pouca (ou nenhuma) experiência culinária
- ☐ Falta de vontade para cozinhar
- ☐ Escassez de alimentos em casa
- ☐ Vontade de um tipo específico de refeição
- ☐ Variedade de escolha
- ☐ Promoções
- ☐ Comodidade
- ☐ Diversidade de métodos de pagamento
- ☐ Segurança perante as novas circunstâncias (COVID-19)
- ☐ Outro: _____

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

Que entraves sente que se impõem quando pretende encomendar uma refeição online? *

Selecione todas as opções com que se identifique

- ☐ Preço da refeição
- ☐ Tempo de espera
- ☐ Carência do método de pagamento preferível
- ☐ Possíveis complicações associadas ao erro humano nos restaurantes
- ☐ Receio perante as novas circunstâncias (COVID-19)
- ☐ Nenhum
- ☐ Outro: _____

Voltar

Enviar

Padrões dos consumidores do mercado online de entrega de refeições ao domicílio

Referências

- Belfo, F. (2020). *Apresentação Resumida e Adaptada do Modelo CRISP-DM*. Coimbra: ISCAC | Coimbra Business School. ISCAC | Coimbra Business School.
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *CRISP-DM 1.0: Step-By-Step Data Mining Guide*. SPSS, CRISP-DM Consortium.