



UNIVERSIDADE CATÓLICA PORTUGUESA

O estudo da correlação dos hábitos alimentares na definição do prémio do seguro

Ana Sofia de Serpa Pinto Pereira Gomes

Católica Porto Business School
Março de 2019



UNIVERSIDADE CATÓLICA PORTUGUESA

O estudo da correlação dos hábitos alimentares na definição do prémio do seguro

Trabalho Final na modalidade de Projeto
apresentado à Universidade Católica Portuguesa
para obtenção do grau de mestre em Gestão de Serviços

por

Ana Sofia de Serpa Pinto Pereira Gomes

sob orientação de
Prof. Doutor Jorge Julião

Católica Porto Business School
Março de 2019

“Não há desenvolvimento físico ou intelectual sem esforço, e esforço significa trabalho.”
Calvin Coolidge

Agradecimentos

A realização deste trabalho requereu dedicação e rigor, bem como o apoio de pessoas que ao longo deste percurso prestaram contributos inestimáveis, e aos quais pretendo agradecer.

Em primeiro lugar, gostaria de agradecer ao meu orientador, Professor Doutor Jorge Julião, pelo seu apoio, cooperação e disponibilidade ao longo de todo o estudo, por me ajudar a desenvolver e concluir um projeto tão desafiador como este.

Em segundo lugar, um sincero obrigado à empresa que permitiu a realização do presente projeto, bem como às pessoas que a compõem e que sempre me motivaram, em especial ao Tiago Osório, pelo desafio que me lançou, pela sua exigência e participação ativa.

Aos meus pais e família, que ao longo da minha vida foram, e são, o meu maior alicerce e que tudo fazem por mim.

Por fim, ao meu namorado e amigos, um apoio incondicional, o meu obrigada por toda a motivação, compreensão e ajuda ao longo deste desafio.

Resumo

Tendo como mote o estudo do desenvolvimento de um novo modelo de negócio, disruptivo pelo modo inovador utilizado para determinar o risco associado a um seguro, o presente trabalho pretende fornecer um contributo para o conhecimento existente sobre o tema. O processo de transformação digital que enfrentamos atualmente apresenta novas oportunidades às empresas que pretendem apostar em novas soluções e na diferenciação nas suas áreas de negócio. Este estudo está, assim, assente na aplicação dos padrões de consumo alimentares por forma a que possam resultar em benefícios para o consumidor, através de descontos associados ao prémio de um seguro de saúde ou de vida. O modelo proposto por esta investigação sugere uma metodologia de classificação do cabaz de consumo alimentar existente no retalhista, capaz de ser aplicado e replicado para todas as transações realizadas. A definição de uma métrica que classifique o cabaz alimentar do cliente, considerando a proporção do produto classificado bem como o agregado familiar, permite a customização do prémio consoante o perfil do cliente. O estudo desenvolvido inerente à classificação das categorias alimentares dos retalhistas apresenta 44% dos produtos como saudáveis. Em contraste, a análise em *Power BI Desktop* de uma amostra aleatória, evidencia que 37% dos produtos obtiveram uma classificação saudável, sendo que considerando a repetição da compra pelos clientes durante o período de análise, o peso destes produtos representa 51% do total.

Palavras-chave: Semáforo Nutricional; *InsurTech*; Padrão de Consumo Alimentar; Seguro Ramo Vida.

Abstract

Taking as its motto the study of the development of a new business model, disruptive by the innovative way used to evaluate the risk associated with an insurance, the present study intends to contribute to the existing knowledge on the subject. The digital transformation process that we are facing today presents new opportunities for companies that want to invest on new solutions and differentiation in their business area. Therefore, this study analyses the impact of the discounts associated with the premium of health or life insurance in consumer's benefits through food consumption patterns.

The proposed model presents a classification methodology of the food consumption basket existing in the retailer which is capable to be applied and replicated for all transactions. Taking into account the proportion of the classified product as well as the household, the definition of a metric that categorizes the client's food basket allows the customization of the premium according to the client's profile.

The study argued in this dissertation settles that 44% of the retailer's food are classified as healthy. In contrast, the analysis of a random sample through Power BI Desktop reveals that 37% of the products obtained a healthy classification. Despite the recurrence of the purchase by the customers during the period of analysis, the weight of these products represents 51% of the total

Keywords: Nutritional Traffic Light; InsurTech; Food Consumption Patterns; Life Insurance.

Índice

Agradecimentos	v
Resumo	vii
Abstract	viii
Índice	ix
Lista de Figuras	xi
Lista de Tabelas	xii
1. Introdução	13
1.1. <i>Motivação Pessoal</i>	13
1.2. <i>Enquadramento</i>	14
1.3. <i>Desafio e objetivos de investigação</i>	16
1.4. <i>Metodologia</i>	18
1.5. <i>Estrutura do trabalho</i>	19
2. Revisão Bibliográfica	20
2.1. <i>Desenvolvimento tecnológico enquanto potenciador da economia</i>	22
2.1.1. <i>Impactos no mercado</i>	23
2.1.2. <i>Impactos nos serviços financeiros</i>	25
2.2. <i>Hábitos de consumo alimentares</i>	28
2.2.1. <i>Hábitos alimentares e o género</i>	30
2.2.2. <i>Hábitos alimentares e o status socioeconómico</i>	30
2.2.3. <i>Hábitos alimentares e a saúde</i>	33
2.3. <i>Estudos semelhantes</i>	34
2.3.1. <i>A relevância dos cartões de fidelização</i>	35
2.3.2. <i>A informação nutricional no rótulo das embalagens</i>	37
2.4. <i>Sumário</i>	40
3. Metodologia	42
3.1. <i>Metodologia de Investigação</i>	42
3.2. <i>Conceção do Modelo-SN</i>	44
3.3. <i>Teste do Modelo-SN</i>	50
3.4. <i>Sumário</i>	54
4. Apresentação e discussão dos resultados	56
4.1. <i>Apresentação dos resultados</i>	56
4.2. <i>Discussão</i>	61
5. Conclusão	65
5.1. <i>Implicações</i>	65
5.2. <i>Forças e limitações do estudo</i>	66
5.3. <i>Desenvolvimentos Futuros</i>	68
Referências Bibliográficas	70

Apêndices	74
<i>Apêndice I – Exemplo do Modelo-SN.....</i>	<i>74</i>
<i>Apêndice II – Ilustração da amostra em Power BI Desktop</i>	<i>75</i>

Lista de Figuras

Figura 1 - Interligações necessárias ao desenvolvimento do novo modelo de seguro	17
Figura 2 – Níveis de agregação de produtos existentes no Retalhista	45
Figura 3 - Relações estabelecidas nas tabelas de análise	52
Figura 4 - Média do Score Final discriminado de acordo com o Modelo-SN .	58
Figura 5 - Média do Score Final discriminado por género	59
Figura 6 - Média do Score Final discriminado pelo Segmento Estilo de Vida	59
Figura 7 – Distribuição dos clusters de acordo com o total de clientes e Score Final.....	60

Lista de Tabelas

Tabela 1 - Modelo de Inovação das InsurTech	28
Tabela 2 - Região de análise dos estudos comparados	32
Tabela 3 – Semáforo Nutricional para Alimentos	38
Tabela 4 – Semáforo Nutricional para Bebidas	38
Tabela 5 – Semáforo Nutricional para Alimentos	39
Tabela 6 – Semáforo Nutricional para Bebidas	39
Tabela 7 – Exemplo da estrutura da base de dados do retalhista	45
Tabela 8 – Exemplo ilustrativo de classificação do Modelo-SN.....	48
Tabela 9 – Distribuição do Modelo-SN por transações e SKU	58
Tabela 10 - Distribuição dos clientes de acordo com os clusters definidos	61
Tabela 11 - Segmentação do Cabaz de consumo cabaz de implicar descontos associados ao prémio	64

1. Introdução

1.1. Motivação Pessoal

Se considerarmos a tecnologia como todas as criações feitas pelo homem ao longo da sua existência, podemos afirmar, decisivamente, que nos últimos anos os avanços têm sido vertiginosos. Não foi só a tecnologia *per si* que promoveu o surgimento de novos produtos e ofertas, foi também a ânsia do homem em procurar novas ideias e soluções, em sonhar e dar liberdade ao seu cérebro para desafiar os dogmas e limitações que os seus antecessores e muitos dos seus pares ainda acreditam. Mas o que realmente deu força a este movimento evolutivo foi o enorme investimento feito à escala global no ensino e formação, estendido a toda a população e não apenas a uma classe de privilegiados. Aprender, sonhar, testar, errar, reformular e acertar. Para ilustrar um pouco melhor a imagem que pretendo transmitir, coloco de seguida uma referência de um ilustre pensador:

“No edifício do pensamento não encontrei nenhuma categoria na qual pousar a cabeça. Em contrapartida, que belo travesseiro é o Caos!” – Emil Cioran

Foi através desta linha de pensamento que surgiu a ideia de avançar com este projeto. No âmbito do meu trabalho atual tive a oportunidade de estudar a envolvente dos seguros, procurando fatores diferenciadores e tendências que pudessem trazer valor acrescentado à oferta que estamos a desenvolver. Sendo que hoje os conceitos *inovação* e *evolução* são cada vez mais utilizados por todos, importa encontrar nichos e focos de atuação onde a perceção do cliente seja positivamente afetada. É neste contexto que surge a ideia de modelar uma solução que permita identificar padrões de consumo do indivíduo e, cruzando-os com modelos de conhecimento, obter um perfil ou *score* que possa depois ser aproveitado para influenciar o custo a pagar pelo utilizador.

1.2. Enquadramento

As seguradoras atualmente presentes no mercado nacional não estão a aproveitar o desenvolvimento visível da tecnologia para melhorar a proposta de valor associada aos seguros que disponibilizam aos consumidores. Um exemplo da resistência à mudança e inovação sentida nas seguradoras portuguesas prende-se com a reduzida oferta de seguros com processos totalmente, ou predominantemente, digitais, numa era onde existe uma nova franja de pessoas que privilegia a utilização de novas tecnologias e que está ávida de novidades e soluções alinhadas com a forma como olham para o mundo. No caso concreto dos seguros de saúde ou de vida, a aplicação de dados não tradicionais existentes sobre o cliente – como hábitos de consumo ou estilo de vida – na determinação mais correta do risco associado a cada consumidor exemplifica uma potencialidade da transformação digital ainda por aproveitar. Gundlach (2017) destaca que, especialmente em indústrias caracterizadas por elevada competição, a inovação é considerada uma atividade crucial para que se possam distinguir da concorrência.

O Grupo onde o presente estudo se realiza gere um portefólio diversificado de negócios nas áreas de retalho, serviços financeiros, tecnologias, centros comerciais e telecomunicações. Mais concretamente, a empresa onde o projeto se desenvolveu é a unidade de negócio que coordena os serviços financeiros do Grupo. Em outubro de 2015, a empresa lançou um cartão de crédito e de fidelização. Adicionalmente, a empresa conta com soluções de financiamento, bem como, através da mediação de seguros, disponibiliza aos consumidores seguros associados à proteção do crédito em cartão, do orçamento familiar, da casa, e proteção em caso de doenças graves.

Tendo presente um dos valores e princípios presentes na abordagem aos negócios do Grupo – inovação – o desenvolvimento do presente projeto visa

desafiar os paradigmas atuais e identificar novas oportunidades de negócio. Deste modo, na procura pela melhoria contínua da proposta de valor, o presente projeto pretende contribuir para a conceção de um novo modelo de seguro, disruptivo pelo modo inovador utilizado para determinar o risco e conseqüente prémio. Reconhecendo as potencialidades associadas ao portefólio de negócios do Grupo – nomeadamente através da ligação ao retalho alimentar e respetivo cartão de fidelização – identificou-se como oportunidade o desenvolvimento de um seguro que considere e analise os dados de consumo alimentares.

Posto isto, surge a seguinte questão de investigação – *De que forma se poderá desenvolver um seguro que beneficie o cliente consoante o seu perfil de consumo?* Deste modo, a definição deste seguro não irá implicar uma penalização para o consumidor, mas antes uma forma de recompensar o cliente pela prática de uma alimentação saudável. Assim, no momento da adesão, será definido um prémio independente dos hábitos de consumo, que poderá ser reduzido, posteriormente, consoante o perfil do cliente. Contudo, para que a atribuição de descontos seja benéfica para a empresa, apenas os clientes com padrões de consumo saudáveis, isto é, com menor risco, serão premiados com uma redução do preço do seguro. Ora, este princípio apresenta uma característica fundamental para o cliente autorizar o acesso aos seus dados de consumo (Coll, 2013).

A definição do tema a estudar parte, então, da identificação de uma lacuna no conhecimento atual, dado que não existem métricas de determinação do risco associado a um seguro de saúde ou de vida através do cruzamento de dados relacionados com os hábitos de consumo alimentar. Esta problemática acontece porque as seguradoras portuguesas utilizam medidas de *scoring* tradicionais – como a idade e hábitos tabagistas, ainda não adaptadas às potencialidades do *big data*. As empresas que evoluem com a Transformação Digital, para além de acompanharem as exigências do mercado, conseguem alcançar maiores

vantagens competitivas face aos seus principais concorrentes (Roedder et al., 2016).

Teece (2010) define modelo de negócio enquanto o modo como a empresa cria valor para o consumidor, refletindo-se no valor pago pelos consumidores e nos lucros para a empresa. Este novo modelo de negócio poderá incluir a definição de segmentos-alvo, e beneficia da dificuldade de replicação do mesmo por concorrentes (Teece, 2010). Como tal, o desenvolvimento de um novo modelo de negócio, para além de beneficiar os consumidores, permite às empresas a criação de vantagens competitivas, traduzidas num aumento de rentabilidade. Esta fonte de rentabilidade poderá ser proveniente i) do aumento do volume de vendas; ii) da redução do risco num médio e longo prazo, associado ao incentivo em adotar comportamentos mais saudáveis; ou iii) de um aumento da fidelização dos clientes aos parceiros, incentivando-os a suportar os custos da atribuição de descontos.

1.3. Desafio e objetivos de investigação

De forma a estudar este tema, definiu-se um objetivo principal de investigação – *contribuir para o desenvolvimento de um novo modelo de negócio de seguros, cuja determinação do risco associado considere os hábitos alimentares dos clientes*. Mais especificamente, este objetivo subdivide-se em dois objetivos complementares: i) *contribuir para a definição de uma metodologia de classificação dos bens alimentares transacionados*, isto é, desenvolver ferramentas de análise que segmentem um determinado produto, avaliando se o mesmo é ou não saudável; e ii) *analisar se a metodologia desenvolvida no primeiro ponto é capaz de ser replicável para o universo das transações e clientes fidelizados*.

A presente questão de investigação, pelo carácter abrangente como foi construída, permite que a mesma seja aplicada por um conjunto amplo de empresas. Contudo, este projeto pretende estudar e desenvolver um novo modelo de seguros que possa ser beneficiado por uma empresa concreta. Desta forma, considera-se relevante apresentar as interligações existentes e essenciais ao desenvolvimento deste modelo de seguro pela empresa. Assim, destacam-se quatro principais *stakeholders*, excluindo o cliente final: i) a marca; ii) a detentora do cartão de crédito e fidelização; iii) a Seguradora; e iv) a empresa do Grupo de retalho alimentar, que detém o programa de fidelização e que permite o acesso ao conhecimento da informação transaccional do cliente no retalhista. De uma forma esquemática, a **Figura 1** apresenta as relações existentes, identificando a cinzento os dois intervenientes que serão analisados neste projeto.

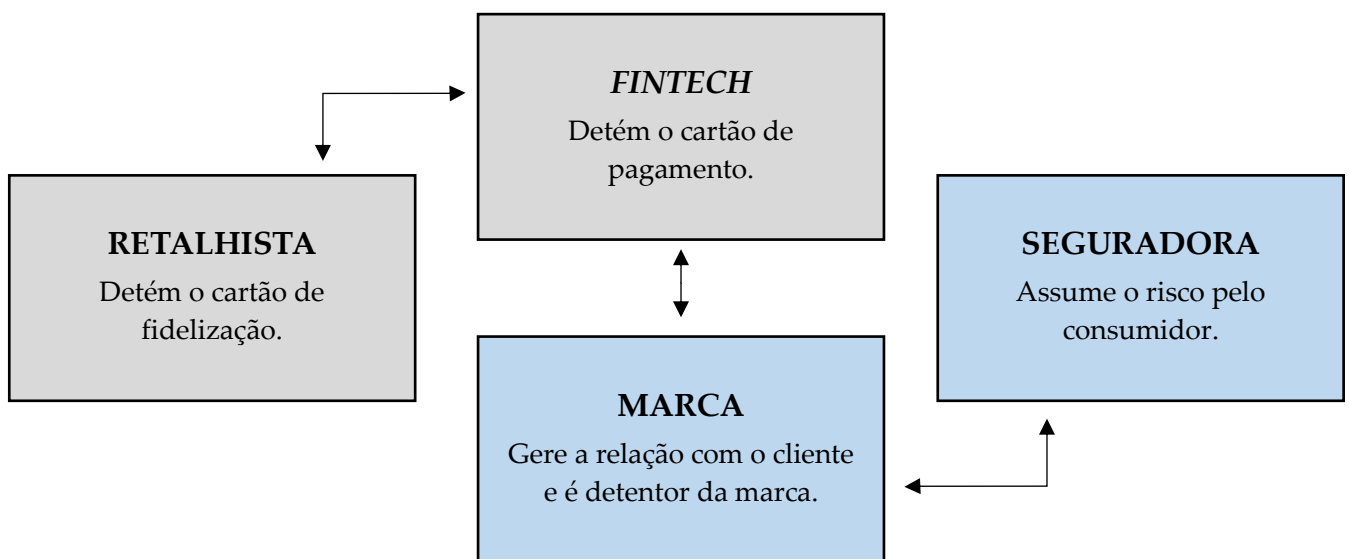


Figura 1 - Interligações necessárias ao desenvolvimento do novo modelo de seguro

Deste modo, esta tese de mestrado irá centrar-se em desenvolver conhecimento capaz de contribuir para a criação de um novo modelo de negócio de seguros, nomeadamente um seguro de vida ou de saúde, cuja determinação do risco associado considere os hábitos alimentares dos clientes através do cartão de fidelização. Considera-se importante realçar que, dado o carácter inovador da

investigação, este trabalho deverá ser compreendido como um ponto de partida para um novo modelo de negócio, procurando fornecer alguma orientação para trabalhos futuros. Como tal, neste estudo não serão abordados modelos de *scoring* de avaliação do risco, nem modelos de *pricing* ou de comunicação da proposta de valor gerada. Do mesmo modo, uma vez que não foram identificados estudos científicos que proponham uma metodologia de aplicação do modelo de classificação dos produtos na classificação do cabaz alimentar de um consumidor, a presente investigação pretende fornecer contributos para uma determinação mais correta desta relação, não estando ainda testada ou comprovada.

1.4. Metodologia

De forma a alcançar os objetivos definidos, foi utilizada uma estratégia de investigação explorativa de natureza qualitativa, no sentido de analisar o conhecimento existente relacionado com o tema. De seguida, foi realizada uma revisão documental, que permitiu a identificação de um modelo de classificação dos produtos pré-embalados, desenvolvido com o objetivo de fornecer aos consumidores uma forma rápida e intuitiva de identificar se determinado produto é mais ou menos saudável, tendo em conta a menor ou maior quantidade de certos nutrientes. Por conseguinte, pretende-se aplicar este princípio num modelo estruturado de acordo com a categoria do produto comprado, permitindo assim a replicação do mesmo. Em terceiro lugar, a aplicação dos resultados numa amostra aleatória será alcançada com recurso à extração de dados transacionais, e respetivas dimensões de produtos e clientes, fornecidos pelo cartão de fidelização. Deste modo, será possível a identificação

de padrões de consumos de clientes, e a criação de segmentos que permitirá a aplicação de descontos associados ao prémio do seguro.

1.5. Estrutura do trabalho

A conceção e estruturação do presente projeto compreende, primeiramente, a inclusão de um capítulo introdutório onde são apresentadas as motivações e desafios identificados, onde é enquadrado o problema a estudar no contexto empresarial, e metodologia de investigação utilizada. Numa segunda fase, aprofundar-se-á a literatura existente sobre o tema, focada no potencial do desenvolvimento tecnológico na economia, no conhecimento existente sobre os hábitos de consumo alimentares, e por fim reunir-se-á práticas semelhantes que possam ser replicadas. Num terceiro momento, explicar-se-á a metodologia adotada e utilizada quer na definição do modelo, como na validação do mesmo. Finalmente, terá lugar a apresentação e discussão dos resultados, terminando o presente trabalho com as conclusões obtidas, principais pontos fortes e limitações, e contribuições futuras.

2. Revisão Bibliográfica

A evolução das tecnologias atuais, acompanhadas pelo avanço geracional, conduziram a novas exigências e desafios no mercado. Por um lado, devido à sua potencialidade, são diversas as áreas que atualmente estão interessadas em aproveitar os meios digitais para estudar os seus clientes. Ao aproveitar os desenvolvimentos tecnológicos para evoluir e conhecer melhor o cliente, as indústrias puderam aperfeiçoar, reformular ou criar novos produtos. Por outro lado, os consumidores têm hoje cada vez mais acesso a um conjunto superior de bens e de conhecimento, o que os torna mais exigentes para com a oferta disponível no mercado. Importa tomar consciência de que as novas gerações procuram soluções diferenciadoras, capazes de oferecer serviços com simplicidade, customização e conveniência.

Nos serviços financeiros, esta nova realidade levou ao surgimento de novas soluções promovidas pela tecnologia – as *FinTech* – que se diferenciam dos Bancos Tradicionais por oferecerem soluções mais simples e disruptivas. Particularmente, um ramo dos serviços financeiros que se tem destacado internacionalmente pelo desenvolvimento de soluções tecnológicas é o ramo segurador, caracterizado pela transferência de risco de um consumidor para a empresa. As *InsurTech* são novas entrantes no mercado que oferecem produtos mais simples, flexíveis e centrados no consumidor, face às seguradoras tradicionais. Estas, apesar de ainda não conseguirem oferecer soluções semelhantes às *InsurTech*, têm começado a desenvolver infraestruturas *data-driven*, que recorrem à utilização de dados relacionados com o risco para determinar, de uma forma mais precisa, o prémio associado.

Ora, é através da identificação da oportunidade de diferenciação que a presente revisão da literatura se desenvolve. Deste modo, estuda-se a criação de

programas de fidelização por parte dos retalhistas, que passam a deter um conjunto de dados do cliente no momento da compra, que permite às empresas conhecerem o seu cliente e poderem segmentar as campanhas de *marketing* (Coll, 2013; Meyer-Waarden, 2008). Como a definição apresentada demonstra, os programas de fidelização posicionam-se como ferramentas valiosas, capazes de serem utilizadas no estudo dos clientes. No estudo em concreto, este conhecimento apresenta o potencial de poder ser aplicado na avaliação da prática de uma dieta saudável e percutível no prémio de um seguro de saúde ou de vida.

Até à data não foram encontrados estudos que apliquem uma relação entre os hábitos de consumo alimentares e o prémio de um seguro de saúde. No entanto, existem diversos estudos que demonstram o impacto dos hábitos de consumo na saúde, bem como estudos que relacionam os padrões alimentares com o género e afluência dos consumidores. Adicionalmente, sendo a preocupação com o estilo de vida e alimentação saudável temas cada vez mais estudados, têm vindo a ser desenvolvidos mecanismos que permitem aos consumidores aferir, de uma forma facilitada, se os alimentos que estão a seleccionar são mais ou menos saudáveis.

O presente capítulo procura em primeiro lugar identificar oportunidades de melhoria que possam ser aproveitadas pelas empresas que comercializam seguros, bem como reunir o conhecimento existente sobre os hábitos de consumo alimentares e a respetiva relação entre o género, status socioeconómico das famílias, e saúde dos indivíduos. No final são apresentados modelos semelhantes de análise dos hábitos de consumo associados a um cartão de fidelização.

2.1. Desenvolvimento tecnológico enquanto potenciador da economia

Ao longo dos últimos 10 anos, temos assistido a uma constante progressão e desenvolvimento das tecnologias digitais (Gölzer & Fritzsche, 2017). Atualmente, são diversas as áreas – desde a ciência, economia e sociologia – que investigam o acesso a grandes quantidades de informação produzida sobre pessoas, recursos e respetivas interações (Boyd & Crawford, 2012). Devido ao desenvolvimento dos computadores pessoais e da *internet*, a produção, partilha, interação e organização de dados passou a ser possível para um conjunto alargado de pessoas, desde os estudantes aos profissionais (Boyd & Crawford, 2012).

Roedder et al. (2016) identificaram três principais implicações do desenvolvimento tecnológico. Em primeiro lugar, cada desenvolvimento tecnológico tem um potencial económico, capaz de ser utilizado em modelos de negócio. No entanto, é fulcral que a sociedade aceite este desenvolvimento, caso contrário não terá resultados económicos. Em segundo lugar, quando existe um desenvolvimento tecnológico, é necessário analisar se a sua aplicação e utilização cumprem as imposições legais. Por último, assistimos a uma constante preocupação com a segurança e proteção de dados – condição básica para a aceitação de novos modelos de negócio.

Face ao desenvolvimento tecnológico atual, e tendo em conta o âmbito em que a presente revisão da literatura se insere, considera-se relevante apresentar dois conceitos que potenciam o desenvolvimento do presente projeto, são eles o *Big Data* e processo de Transformação Digital. De seguida, serão identificados os impactos nos modelos de negócio dos Serviços Financeiros, especificamente no ramo segurador.

2.1.1. Impactos no mercado

Com a evolução dos meios analíticos, tornou-se possível tratar um conjunto mais robusto de dados para criar processos de análise, previsão e decisão automatizada (Gölzer & Fritzsche, 2017; Roedder et al., 2016).

De acordo com Boyd & Crawford (2012), o termo *Big Data* está associado à capacidade de pesquisar, agregar e cruzar grandes quantidades de dados. Uma vez que atualmente é possível realizar análises robustas de dados em computadores com um *software* standard, a definição de *Big Data* não deverá cingir-se à utilização de grandes quantidades de dados que necessitam de processadores robustos para os estudar. Assim, o termo *Big Data* é comumente conhecido pelo cruzamento de três atributos, os 3V: i) volume de dados tratados; ii) velocidade associada ao processamento e transmissão dos dados – em *batch*, *near time*, *real time* ou *streams*; e iii) variedade, incluindo dados estruturados, não estruturados e semiestruturados (Ahmad et al., 2019; Gölzer & Fritzsche, 2017; Russom & Org, 2011). Gölzer & Fritzsche (2017) defendem que a significância dos dados utilizados não depende apenas da sua origem, mas também da metodologia utilizada no seu tratamento. Como tal, atribuem duas novas características aos 3V: iv) valor da aplicação; e v) veracidade da análise.

Para Boyd & Crawford (2012), o fenómeno *Big Data* resulta i) da interação entre a maximização do poder da computação e da precisão algorítmica a fim de recolher, analisar, interligar e comparar conjuntos de dados; ii) da análise de grandes conjuntos de dados para identificar padrões e aplicar conhecimento; e iii) da crença de que esta informação oferece uma forma superior de inteligência e conhecimento.

De acordo com Hess et al. (2016), integrar e explorar as novas tecnologias digitais corresponde a um dos maiores desafios que as empresas enfrentam atualmente. Quando falamos em Transformação Digital, não nos referimos

apenas a produtos, processos, canais de venda e cadeia de valor, mas sim a todo o modelo de negócio e proposta de valor, pelo que o estudo da sua abordagem deverá ser feito de forma estruturada, sob a pena de poder fracassar (Hess et al., 2016).

Assim, com a Transformação Digital, surgiram novos modelos de negócios: como o caso da *Uber* – empresa de transporte privado com serviços solicitados por via eletrónica e cujo preço praticado é determinado de acordo com o cruzamento entre a oferta de motoristas e a procura por viagens (Roedder et al., 2016); ou do *AirBnB* – plataforma que permite aos proprietários o arrendamento de curta duração a clientes interessados (Dawson, 2016). Por outro lado, a Transformação Digital potenciou a revisão dos modelos de negócio atuais, levando a uma melhoria dos processos internos existentes, através de serviços complementares de informação e suporte aos consumidores; e paralelamente levou ainda a um aumento de eficiência, devido ao uso de métodos preditivos e análise de dados capazes de tomar melhores decisões e/ou decisões autónomas (Roedder et al., 2016). São exemplo de empresas que beneficiaram destas melhorias a *Walmart* e a *Zara* (Dawson, 2016). Por fim, a transformação digital mudou, também, a forma como os consumidores valorizam um determinado bem e serviço. Considere-se por exemplo a música, onde anteriormente os consumidores compravam vários *CD's* para ouvir as canções que gostavam e agora pagam uma única mensalidade de *Spotify* para terem acesso a uma miríade de artistas, géneros e temas inacessíveis anteriormente (Dawson, 2016).

A natureza disruptiva da Transformação Digital requer e impulsiona estratégias, processos, estruturas organizacionais, produtos e serviços em diferentes setores, a ritmos diferentes consoante o setor em questão (Stoekli et al., 2018). As empresas que evoluem com a Transformação Digital, para além de acompanharem as exigências do mercado, conseguem alcançar maiores vantagens competitivas face aos seus principais concorrentes (Roedder et al.,

2016). Como tal, esta evolução tornou-se prioritária na estratégia de diversas empresas como forma de ganhar vantagem competitiva no mercado (Hess et al., 2016). Gölzer & Fritzsche (2017) defendem que, para permanecerem competitivas no mercado, as empresas deverão assegurar que os seus modelos operativos são flexíveis.

2.1.2. Impactos nos serviços financeiros

Em primeiro lugar, importa realçar que as soluções provenientes do desenvolvimento tecnológico nos Serviços Financeiros são bastante recentes, pelo que o conhecimento académico estruturado existente sobre as mesmas é ainda reduzido e mais escasso no caso concreto do ramo segurador (Stoekli et al., 2018).

Relativamente às Financeiras, a transformação digital tem vindo a impactar os bancos líderes tradicionais, que procuram desenvolver soluções digitais. Ao mesmo tempo as *FinTech* desafiam o mercado com soluções simples, disruptivas e orientadas para o consumidor. O termo *FinTech* é utilizado para referir o uso de tecnologia na oferta de soluções financeiras. Deste modo, as *FinTech* são soluções de tecnologia da informação inovadoras, utilizadas por fornecedores de serviços financeiros ou novos entrantes para desenvolver modelos de negócios no setor de serviços financeiros (Stoekli et al., 2018). Presentemente, são diversas as *FinTech* que desafiam o mercado, desde empresas que recorrem a mecanismos *peer-to-peer* para eliminar as taxas que os bancos cobram para trocar dinheiro de uma moeda para outra (*TransferWise*), às alternativas bancárias digitais como bancos digitais (*Monzo*, *N26* e *Starling Bank*), os serviços bancários digitais (*Revolut*), contas bancárias (*Metro Bank*), entre outros (Santos, 2018).

No caso das Seguradoras, a natureza disruptiva do desenvolvimento digital faz, por um lado, com que as grandes seguradoras tradicionais se vejam forçadas a rever o seu modelo de negócio para que possam manter a liderança no mercado,

especialmente nos canais digitais; por outro lado, possibilitou o surgimento das *InsurTech*, que desafiam o mercado ao aproveitarem o potencial das tecnologias da informação como inovação no ramo segurador (Gundlach, 2017; Stoeckli et al., 2018). As *InsurTech* surgem, assim, como o ramo específico dos seguros das *FinTech*, que emergem no mercado por desenvolverem produtos mais simples, flexíveis e centrados no consumidor, alcançando um público-alvo mais alargado (Stoeckli et al., 2018). Assim, as *InsurTech* destacam-se no mercado em modelos de negócio como i) plataformas *peer-to-peer* (*Lemonade*, *Friendsurance* e *Bought by Many*); ii) plataformas impulsionadas pelo desenvolvimento de dispositivos conectados (*Metromile*); iii) mecanismos *on demand* (*Trov*); iv) plataformas de *marketplace* (*SimplyInsured*), entre outros (Gundlach, 2017).

De uma forma simples, o negócio dos seguros consiste, essencialmente, na transferência do risco de um consumidor para um fornecedor de seguros que, com base no risco avaliado, cobra uma respetiva quantidade de dinheiro. Tendo em conta esta premissa, o desenvolvimento da tecnologia de informação e a existência de dados permitem uma medição do risco mais exata (Stoeckli et al., 2018). A exploração destas tecnologias – como o caso de *wearables* que acompanham, monitorizam e aperfeiçoam o risco associado a um seguro de saúde – ainda está pouco aprofundada (Stoeckli et al., 2018). Considera-se relevante realçar o seguro de vida da seguradora *John Hancock* com a *Vitality*¹, uma solução inovadora que atribui pontos pela prática de um estilo de vida saudável. Para tal, quando o cliente adere ao seguro recebe um *smartwatch*, e a quantia associada à mensalidade paga pelo relógio é determinada pela quantidade de exercício físico realizado pelo consumidor durante o mês. Para desenvolver este seguro, a seguradora desenvolveu uma parceria com o

¹ <https://www.johnhancockinsurance.com/vitality-program.html> (acedido a 12.11.2018)

programa *Vitality*, que já se encontra subscrito por mais de 8 000 000 de consumidores, em 19 países².

Com base na análise realizada a diversas *InsurTech*, área de atuação e respetivas fontes de vantagem competitiva e criação de valor no mercado, Stoeckli et al. (2018) desenvolveram um Modelo de Inovação das *InsurTech*, que consta representado na Tabela 1.

Atividade Primária	Segmento	Capacidade de Transformação Emergente
Operações de infraestrutura	Infraestrutura digital	Desenvolvimento de infraestruturas digitais, reduzindo custos de atendimento, enquanto aumenta a diferenciação com a criação de serviços de valor acrescentado e novos pontos de contacto;
	Infraestrutura <i>data-driven</i>	Exploração de dados existentes relacionados com o risco para uma determinação mais precisa e correta do risco, impactando a subscrição;
		Exploração dos dados para gestão de sinistros, capaz de reduzir os custos de tratamento de sinistros ou deteção de fraude.
Provisionamento de serviços	Serviços digitais	Oferta de serviços relacionados nos canais digitais;
		Complementar a oferta do seguro com serviços de prevenção e reparação (e.g. seguro de casa que inclui câmara de vigilância), aumentando o valor para o consumidor;
		Integração do seguro com produtos relacionados para aumentar a frequência de contacto entre seguradora e consumidor.
	Desenvolvimento de serviços	Capacidade de perceber necessidades do consumidor para desenvolver ofertas diferenciadas e alcançar vantagem competitiva;

² <https://www.vitalitygroup.com/vitalityapplestudy/> (acedido a 26.02.2019)

		Ajustar produto de acordo com natureza do risco seguro;
		Desenvolvimento de seguros que cobrem novos riscos (não contemplados antes da Transformação Digital);
		Oferta de um preço dinâmico e ajustado ao risco, capaz de oferecer preços e serviços competitivos.
Promoção em rede e gestão de contrato	Promoção a uma rede de clientes	Desenhar e decidir canais de distribuição digitais de seguros;
		Selecionar e adquirir os clientes certos através do aproveitamento das oportunidades do marketing digital;
		Agir como um mediador digital, que vende seguros com determinada independência da seguradora.
	Parcerias em rede	Criação de parcerias estratégicas para criar vantagem competitiva.

Tabela 1 - Modelo de Inovação das *InsurTech*

Fonte: Tabela adaptada de Stoeckli et al. (2018)

As vantagens competitivas acima apresentadas são características das *InsurTech* e ainda pouco aproveitadas pelas seguradoras tradicionais devido à infraestrutura existente nas segundas. No entanto, tem-se verificado que algumas seguradoras já estão a tomar os primeiros passos para desenvolver infraestruturas *data-driven* (Stoeckli et al., 2018).

2.2. Hábitos de consumo alimentares

A Organização Mundial de Saúde define estilo de vida como a forma de viver baseada em padrões ou comportamentos identificáveis, que são determinados pela interação entre características individuais e pessoais, relações sociais e fatores socioeconómicos e ambientais (Costa-Tutusaus & Guerra-Balic, 2016). Tendo em conta esta definição, os autores definiram um conjunto de 5

componentes para medir o estilo de vida de um indivíduo, são eles i) dieta; ii) atividade física; iii) abuso de substâncias; iv) uso racional da tecnologia; e v) higiene. Para a realização do presente estudo, iremos focar-nos na primeira componente: dieta.

Arechavala et al. (2016) basearam-se na Pirâmide de Recomendação Alimentar – pirâmide que representa graficamente a frequência recomendada do consumo de alimentos de acordo com os seus nutrientes para caracterizar uma dieta saudável. Como tal, uma dieta saudável deve incluir o consumo diário de frutas e legumes, laticínios, cereais e batatas; o consumo intercalado de carne, peixe e ovos; e o consumo ocasional de salsichas, doces e produtos manufacturados (Arechavala et al., 2016). Em contraste, uma dieta pouco saudável consiste no consumo de alimentos ricos em açúcar e gordura, com baixos níveis de proteína, gorduras monoinsaturadas (Voráčková et al., 2016).

Mais do que em qualquer outra altura da evolução humana, os consumidores, para além de terem acesso à informação nutricional dos produtos, têm atenção à mesma (Dunford et al., 2017). Com base na informação nutricional existente, bem como no conhecimento sobre a composição de uma dieta equilibrada, é então, possível analisar se os hábitos de consumo de um indivíduo são, ou não, saudáveis.

Os padrões de consumo alimentares poderão apresentar diferentes comportamentos de acordo com diversas variáveis socioeconómicas. Para a realização deste estudo será analisada, numa primeira fase, a relação entre o género e os hábitos alimentares. Numa segunda fase, será analisada a relação entre a dieta e o status socioeconómico e afluência das famílias. Por último, importa compreender o impacto de uma dieta saudável, ou por contraste de uma dieta pouco saudável.

2.2.1. Hábitos alimentares e o género

A variável género representa, por um lado, uma variável a considerar na avaliação do risco associado a um seguro de saúde ou de vida; e por outro lado, uma variável já estudada na comparação dos hábitos alimentares. Como tal, o presente subcapítulo apresenta as conclusões inferidas por diversos autores relativas aos comportamentos alimentares entre o género feminino e masculino.

Ao analisar o consumo de frutas e legumes – alimentos cruciais para a adoção de uma alimentação saudável – concluiu-se que estes alimentos eram consumidos com maior frequência por raparigas do que por rapazes (Arechavala et al., 2016; Conner et al., 2017; Holubcikova et al., 2016). Paralelamente, o consumo de refrigerantes e *fast food* é superior no género masculino (Arechavala et al., 2016; Conner et al., 2017; Holubcikova et al., 2016). No que diz respeito ao consumo de doces, as conclusões dos autores foram contraditórias – Conner et al. (2017) atribuíram um consumo maior às mulheres, e Arechavala et al. (2016) identificaram um consumo de doces mais elevado nos homens. Por último, globalmente, as mulheres têm um consumo alimentar melhor do que os homens (Costa-Tutusaus & Guerra-Balic, 2016; Holubcikova et al., 2016), o que faz com que o excesso de peso seja maior nos homens (Arechavala et al., 2016). Holubcikova et al. (2016) destacam que o facto de as mulheres praticarem uma alimentação mais saudável do que os homens se deve à preocupação destas com a saúde.

2.2.2. Hábitos alimentares e o status socioeconómico

O conhecimento dos consumidores sobre a nutrição pode influenciar a disponibilidade e consumo de alimentos saudáveis (Holubcikova et al., 2016; Voráčová et al., 2016). No entanto, o preço dos alimentos impacta o consumo dos mesmos. Deste modo, o custo e a disponibilidade dos alimentos pode ser uma barreira para a dieta alimentar das pessoas, em especial daquelas com

rendimentos mais baixos (Arechavala et al., 2016; Costa-Tutusaus & Guerra-Balic, 2016; Voráčová et al., 2016). Os alimentos saudáveis são, por norma, mais caros e muitas vezes não são acessíveis às famílias com rendimentos mais baixos, aumentando as disparidades no acesso a dietas saudáveis (Voráčová et al., 2016). Adicionalmente, os autores destacam a disparidade no peso que uma alimentação saudável e sustentável tem para uma família de baixo rendimento em comparação com uma família de alto rendimento: enquanto a primeira gasta cerca de 48% do seu rendimento semanal, a segunda gasta apenas 9%. Para a análise da revisão da literatura importa, então, analisar a influência do rendimento nos hábitos de consumo.

Diversos estudos comprovam que o status socioeconómico tem influência no desenvolvimento de hábitos alimentares saudáveis em criança, o que tende a refletir-se nos hábitos adotados em adultos (Voráčová et al., 2016). A classificação do status socioeconómico de uma família pode ser medida através do rendimento, educação e profissão do agregado familiar (Voráčová et al., 2016). De forma complementar, e para suprir a ausência de conhecimento das variáveis acima apresentadas em estudos realizados em escolas, diversos autores recorrem a um segundo indicador – escala de afluência familiar (Arechavala et al., 2016; Voráčová et al., 2016). Este indicador mede a afluência material das famílias e é calculado com base na disponibilidade de automóvel e número de computadores em casa, disponibilidade quarto próprio e número de férias ou viagens por ano (Arechavala et al., 2016; Voráčová et al., 2016). Para uma interpretação dos resultados mais robusta, e dada a interligação das duas variáveis, as inferências abaixo incluem análises feitas tendo em conta quer o status socioeconómico como a escala de afluência familiar e são apresentadas como uma só.

Ao analisar o consumo de alimentos específicos como carne vermelha, ovos e *fast food*, é possível identificar diferenças relacionadas com a afluência das famílias (Arechavala et al., 2016). Como tal, famílias menos afluentes apresentam

consumos de carne vermelha e ovos mais equilibrada, mas um consumo de *fast food* mais elevado (Arechavala et al., 2016). Por outro lado, o consumo de refrigerantes, doces e pequeno almoço durante a semana não é variável (Voráčová et al., 2016). No que diz respeito ao impacto do consumo de frutas e vegetais, existe alguma disparidade nos resultados encontrados pelos autores. Estas diferenças podem, contudo, dever-se ao local onde o estudo foi realizado, uma vez que a abundância dos alimentos pode variar de acordo com a geografia, e a procura por determinado produto ser, ou não elástica. Deste modo, considera-se relevante apresentar a região de análise de cada estudo comparado, que se encontra apresentada na Tabela 2. Assim, enquanto Thorndike et al. (2017) e Voráčová et al. (2016) identificaram um consumo mais elevado destes alimentos nos grupos mais afluentes, Arechavala et al. (2016) não verificaram diferenças consideráveis no consumo de frutas e vegetais. A inexistência de diferenças associadas ao consumo de frutas e legumes deve-se, maioritariamente, à disponibilidade e acessibilidade dos produtos – visto que no Sul da Europa estes produtos são abundantes e conseqüentemente baratos (Arechavala et al., 2016).

Autores	Região de Análise
Arechavala et al. (2016)	Espanha
Voráčová et al. (2016)	República Checa
Thorndike et al. (2017)	Estados Unidos da América
Holubcikova et al. (2016)	Eslováquia

Tabela 2 - Região de análise dos estudos comparados

Apesar de não identificarem variâncias, Arechavala et al. (2016) assumem que a qualidade e variedade dos frutos e legumes consumidos podem variar consoante a afluência das famílias – enquanto famílias menos afluentes optam por frutos e legumes mais baratos, e por isso, adotam uma alimentação menos variada, as famílias mais afluentes tendem a consumir alimentos mais variados e

dispendiosos. Voráčová et al. (2016) destacam a implementação de medidas públicas que fomentem o consumo de frutas e legumes como forma de reduzir a disparidade entre grupos socioeconômicos. Para além da disponibilidade e preço dos produtos representarem barreiras a uma dieta saudável, Thorndike et al. (2017) referem que o tempo de preparação também condiciona a decisão de compra de frutos e legumes.

Globalmente, as famílias mais afluentes tendem a ser mais preocupadas com a alimentação (Arechavala et al., 2016; Voráčová et al., 2016), com especial destaque para as mulheres (Arechavala et al., 2016). Ao contrário das evidências acima apresentadas, Holubcikova et al. (2016) não identificam diferenças relacionadas com a afluência das famílias.

Por último, importa realçar que pessoas com maior nível de escolaridade tendem a responder ter melhores hábitos alimentares do que realmente têm, o que pode aumentar o *gap* dos resultados entre os consumidores mais afluentes e menos afluentes (Voráčová et al., 2016).

2.2.3. Hábitos alimentares e a saúde

Costa-Tutusaus & Guerra-Balic (2016) afirmam que existe uma clara relação entre o estilo de vida e a saúde dos indivíduos e que, dessa forma é possível caracterizar estilos de vida mais ou menos saudáveis. A prática de hábitos alimentares pouco saudáveis na adolescência leva a uma variedade de problemas e desordens de saúde (Holubcikova et al., 2016). À semelhança do que já foi identificado no capítulo 2.2., o consumo de frutos e vegetais é uma componente essencial para uma dieta saudável. Conner et al. (2017) e Thorndike et al. (2017) afirmam que tal está relacionado com uma maior longevidade e maior saúde física – como por exemplo, menor probabilidade de surgimento de doenças cardiovasculares e mental – associado a menores índices de depressão.

Arechavala et al. (2016) destacam que o excesso de peso ou obesidade aumenta o risco de vir a sofrer doenças cardiovasculares. Indivíduos com hábitos alimentares pouco saudáveis têm maior risco de sofrer de excesso de peso e fadiga, o que pode levar a um maior risco de doenças crônicas (Holubcikova et al., 2016; Voráčová et al., 2016). A Organização Mundial da Saúde afirma que uma dieta pouco saudável é um dos principais fatores de risco para uma série de doenças crônicas, e reconhece 6 principais doenças relacionadas com a dieta: a obesidade, hipertensão e doenças cardiovasculares, cancro, diabetes Tipo II, doenças dentárias e osteoporose³.

Ora, uma vez reconhecida a relação entre os comportamentos alimentares e a saúde dos indivíduos, a adoção de políticas públicas de saúde que visem a prática dietas mais saudáveis são determinantes, pelo que devem ser abordadas por meio de ações de prevenção (Egnell et al., 2018). De entre diversas medidas de prevenção adotadas, destaca-se o desenvolvimento rótulos nutricionais, isto é, inclusão de informação nutricional na frente das embalagens dos produtos pré-embalados (Dunford et al., 2017; Egnell et al., 2018). Este método será aprofundado mais à frente, no capítulo dedicado à inclusão de informação nutricional nos rótulos das embalagens.

2.3. Estudos semelhantes

Dado o caráter inovador deste projeto, não foram identificadas investigações de cariz idêntico, isto é, investigações que proponham a classificação de um cabaz de consumo alimentar associado à promoção de um seguro de saúde ou de vida. Deste modo, a revisão da literatura procurou estudar artigos que identifiquem a

³ <https://www.who.int/dietphysicalactivity/publications/trs916/en/> (acedido a 29.10.2018)

potencialidade do projeto, bem como artigos que permitam a exploração do mesmo.

2.3.1. A relevância dos cartões de fidelização

As relações com os clientes têm sido cada vez mais estudadas na literatura de marketing acadêmico e profissional (Meyer-Waarden, 2008). Para maximizar a interação com os clientes bem como a sua retenção, muitas empresas recorrem a instrumentos de *marketing* e programas de fidelização ou de recompensas (Meyer-Waarden, 2008). Define-se Programa de Fidelização como um sistema integrado de ações de *marketing* que visa tornar os clientes mais leais, através do desenvolvimento de relações personalizadas com os mesmos (Meyer-Waarden, 2008).

Há mais de uma década que os grandes retalhistas recolhem volumes de dados – como a identidade do consumidor, data e hora da compra e a listagem de produtos comprada – com base em transações diárias associadas aos programas de fidelização (Coll, 2013). Como incentivo à criação de um cartão de fidelização, mas também para criar custos de mudança, prolongar o relacionamento e incentivar as possibilidades de *cross-selling* e/ou *up-selling*, os programas de fidelização desenvolvem o sistema de recompensas baseado na geração de um comportamento de compra repetido – como por exemplo, as milhas ou os pontos ganhos normalmente não são transferíveis para outra empresa (Coll, 2013; Meyer-Waarden, 2008).

De acordo com (Meyer-Waarden, 2008), os consumidores associados a um programa de fidelização identificam-se mais fortemente com a empresa, o que é particularmente benéfico para indústrias de consumo frequente, com baixa diferenciação nos fornecedores. Associado ao desenvolvimento digital, e apoiado em técnicas *data mining* – métodos estatísticos e computadorizados usados para produzir conhecimento a partir de bases de dados – os retalhistas são agora

capazes de aproveitar os dados recolhidos para conhecer melhor o cliente e personalizar e segmentar as suas comunicações (Coll, 2013). Um exemplo de uma aplicação do *data mining* no setor do retalho é a sugestão de produtos adequados ao consumidor, de acordo com a idade, estado na vida, estilo de vida, entre outros.

Coll (2013) identificou como uma capacidade do desenvolvimento do *data mining* a transmissão de dados de consumo alimentares para uma seguradora. De realçar que, até à data, não tinham sido desenvolvidos estudos que permitissem a colaboração de programas de fidelização e seguradoras. Para estudar esta oportunidade, Coll (2013) realizou entrevistas a consumidores. Quando inquirida uma jovem de 22 anos sobre a possibilidade de uma seguradora poder recusar a sua adesão a um seguro de saúde com base no seu perfil de consumo, a mesma não achou a análise injusta. Em contraste, ao entrevistar uma pessoa reformada de 65 anos, a mesma afirmou não teria interesse em aderir a esta solução, se a mesma fosse desenvolvida com o único intuito de beneficiar a empresa. Pelo contrário, se este desenvolvimento procurar melhorar os estilos de vida da população, a sua opinião alteraria. Como tal, a perceção, por parte dos consumidores, de que a utilização dos dados é legítima e benéfica para o cliente são fatores fulcrais para a aceitação deste produto. Adicionalmente, a preocupação com a proteção dos dados pessoais é uma constante.

Coll (2013) acrescenta que um projeto como este pode ser alcançado se existirem duas condições: i) se os clientes puderem escolher o fornecimento ou não dos dados; ii) se for atribuído algum incentivo ou redução dos prémios. Caso contrário, apenas os consumidores que considerarem ter hábitos alimentares saudáveis estarão interessados em aderir.

2.3.2. A informação nutricional no rótulo das embalagens

Conforme referido em capítulos anteriores, os rótulos nutricionais na frente das embalagens foram desenvolvidos no sentido de promover hábitos alimentares mais saudáveis e, como tal, ajudar os consumidores a tomar decisões de consumo mais saudáveis no momento da compra (Egnell et al., 2018). O desenvolvimento destes rótulos, para além de impactar os consumidores, encoraja os produtores a reformularem a composição dos seus produtos, para que possam apresentar informações mais saudáveis e positivas (Dunford et al., 2017; Egnell et al., 2018).

Egnell et al. (2018) alertam que, para que esta metodologia seja útil, é necessário que os consumidores consigam compreender a informação fornecida. Como tal, ao longo da última década foram desenvolvidos diversos modelos de apresentação da informação nutricional, que podem ser enquadrados em três categorias: i) rótulos numéricos, incorporação do valor nutricional do produto na frente da embalagem; ii) rótulos com códigos de cores (Semáforo Nutricional) onde cada cor tem associada uma quantidade do respetivo nutriente (vermelho para uma quantidade elevada; amarelo para uma quantidade moderada; verde para uma quantidade reduzida); e iii) rótulos de alerta, com sinais que avisam o cliente quando um produto inclui uma quantidade excessiva para a saúde de um nutriente.

Os sistemas de apresentação da informação nutricional poderão, assim, ajudar a que os consumidores se tornem mais conscientes da prática de uma alimentação saudável. Contudo, estes modelos foram desenvolvidos para classificar apenas produtos pré-embalados, não contemplando soluções para produtos frescos.

Apesar de os resultados serem positivos independentemente do modelo aplicado, o segundo modelo apresenta diferenças mais significativas do que os restantes, visto que a utilização de cores é mais apelativa e a utilização das cores de um semáforo são compreendidas pela comunidade de forma rápida e simples

(Dunford et al., 2017; Egnell et al., 2018). Assim, o Semáforo Nutricional classifica a quantidade de lípidos, saturados, açúcares, e sal existentes em cada produto. Em 2007, a Food Standards Agency (2007) publicou um dos primeiros modelos de classificação de acordo com este sistema, que se encontra apresentado na Tabela 3 e na Tabela 4. Dez anos depois, e baseados no semáforo nutricional definido pelo Departamento da Saúde do Reino Unido, Dunford et al. (2017) analisaram um conjunto de 290 000 produtos existentes nos Estados Unidos da América, de acordo com a Tabela 5 e a Tabela 6.

	Verde (baixo)	Amarelo (médio)	Vermelho (elevado)	
Lípidos	≤ 3,0 g / 100 g	>3,0 e ≤ 20,0 g / 100 g	> 20,0 g / 100 g	> 21 g / porção
Saturados	≤ 1,5 g / 100 g	> 1,5 e ≤ 5,0 g / 100 g	> 5,0 g / 100 g	> 6,0 g / porção
Açúcares	≤ 5,0 g / 100 g	> 5,0 e ≤ 15,0 g / 100 g	> 15,0 g / 100 g	> 18,0 g / porção
Sal	≤ 0,3 g / 100 g	> 0,3 e ≤ 1,5 g / 100 g	> 1,5 g / 100 g	> 2,4 g / porção

Tabela 3 – Semáforo Nutricional para Alimentos

Fonte: Food Standards Agency, 2007

	Verde (baixo)	Amarelo (médio)	Vermelho (elevado)
Lípidos	≤ 1,5 g / 100 ml	> 1,5 e ≤ 10,0 g / 100 ml	> 10,0 g / 100 ml
Saturados	≤ 0,75 g / 100 ml	> 0,75 e ≤ 2,5 g / 100 ml	> 2,5 g / 100 ml
Açúcares	≤ 2,5 g / 100 ml	> 2,5 e ≤ 7,5 g / 100 ml	> 7,5 g / 100 ml
Sal	≤ 0,3 g / 100 ml	> 0,3 e ≤ 1,5 g / 100 ml	> 1,5 g / 100 ml

Tabela 4 – Semáforo Nutricional para Bebidas

Fonte: Food Standards Agency, 2007

	Verde (baixo)	Amarelo (médio)	Vermelho (elevado)	
Lípidos	≤ 3,0 g / 100 g	> 3,0 e ≤ 17,5 g / 100 g	> 17,5 g / 100 g	> 21 g / porção
Saturados	≤ 1,5 g / 100 g	> 1,5 e ≤ 5,0g / 100 g	> 5,0 g / 100 g	> 6,0 g / porção
Açúcares	≤ 5,0 g / 100 g	> 5,0 e ≤ 22,5 g / 100 g	> 22,5 g / 100 g	> 27 g / porção
Sal	≤ 0,12 g / 100 g	> 0,12 e ≤ 0,6 g / 100 g	> 0,6 g / 100 g	> 0,72 g / porção

Tabela 5 – Semáforo Nutricional para Alimentos

Fonte: Dunford et al. (2017)

	Verde (baixo)	Amarelo (médio)	Vermelho (elevado)	
Lípidos	≤ 1,5 g / 100 ml	> 1,5 e ≤ 8,75 g / 100 ml	> 8,75 g / 100 ml	> 10,5 g / porção
Saturados	≤ 0,75 g / 100 ml	> 0,75 e ≤ 2,5 g / 100 ml	> 2,5 g / 100 ml	> 3,0 g / porção
Açúcares	≤ 2,5 g / 100 ml	> 2,5 e ≤ 11,25 g / 100 ml	> 11,25 g / 100 ml	> 13,5 g / porção
Sal	≤ 0,12 g / 100 ml	> 0,12 e ≤ 0,6 g / 100 ml	> 0,3 g / 100 ml	> 0,4 g / porção

Tabela 6 – Semáforo Nutricional para Bebidas

Fonte: Dunford et al. (2017)

Assim, para classificar um determinado produto, deverá ser comparada a quantidade nutricional do mesmo, numa referência de 100 g ou 100 ml, consoante o mesmo seja um alimento ou bebida, respetivamente. No caso de produtos com porções superiores a 250 gramas, a análise deverá ser feita analisando a quantidade nutricional do produto por porção e, caso ultrapasse o limite identificado na respetiva coluna, o nutriente deve receber uma classificação automaticamente vermelha (Food Standards Agency, 2007).

Para analisar os hábitos alimentares da população americana, Dunford et al. (2017) atribuíram uma ponderação de 1 ponto para cada nutriente classificado a verde; 2 pontos para nutrientes classificados amarelo; e 3 pontos a nutrientes classificados a vermelho, o que se traduz numa ponderação global que varia entre 4 e 12 pontos. Para uma ponderação global entre 4 e 6 pontos, os produtos são

classificados como saudáveis. Ao analisar a classificação dada pelo modelo aos nutrientes, mais de 40% dos produtos existentes nos Estados Unidos da América que foram analisados obtiveram classificações vermelhas. Ao analisar a classificação global dos nutrientes, apenas 30,1% dos produtos avaliados foram classificados como saudáveis (Dunford et al., 2017). Por último, os autores realçam que uma limitação do método utilizado é proveniente do mesmo apenas ter em conta a análise dos nutrientes em vez de se basear num padrão alimentar (Dunford et al., 2017).

2.4. Sumário

Ao subscrever um seguro de automóvel, que entidade terá informação mais precisa e atualizada sobre a localização do automóvel e do estilo de condução do detentor do seguro? Isto é, se o mesmo se encontra estacionado numa garagem, se passa por uma localização mais perigosa, se conduz a uma velocidade excessiva face à estrada onde se encontra, ou se fala ao telemóvel enquanto conduz. Será a seguradora? A companhia fornecedora do carro? Ou será a empresa detentora do *smartphone* ou *smartwatch* que o tomador do seguro possui?

Esta experiência demonstra um dos desenvolvimentos tecnológicos associados à indústria seguradora, e levanta um conjunto de oportunidades: para os consumidores, que passam a esperar uma diferenciação no prémio do seu seguro, adaptado ao seu próprio perfil; e para as indústrias digitais, que, devido às suas capacidades, se tornam capazes de oferecer propostas de valor superiores face às indústrias tradicionais (Dawson, 2016).

Assim, a simplificação dos processos e o desenvolvimento de modelos de decisão autónomos, promovido pelo *big data* e ferramentas de *advanced analytics*, permitiram às empresas melhorar os seus serviços. No setor dos serviços

financeiros, percebeu-se que esta área já se encontra a ser explorada, no entanto por um nicho específico de empresas. As financeiras e seguradoras mais tradicionais, por deterem infraestruturas mais rígidas enfrentam maiores desafios ao tentarem adaptar os seus processos para estes novos modelos digitais.

Uma forma de beneficiar dos dados existentes sobre o cliente está associada aos programas de fidelização e ao desenvolvimento de ferramentas de *analytics* que permitam a geração de conhecimento importante sobre o cliente. Ora, no caso de um retalhista alimentar, a existência de um cartão de fidelização permite a identificação de padrões de consumo alimentar dos vários clientes. O tratamento desta informação pode ser aplicado, de forma benéfica, em seguros de saúde ou de vida, dada a relação existente entre a prática de uma dieta saudável e a menor ou maior probabilidade de ocorrência de doenças.

Contudo, apesar de a relação entre a ponderação dos hábitos de consumo alimentares e o risco associado a um seguro de saúde já ter sido estabelecida por Coll (2013), não se encontra desenvolvida, até à data, nenhuma metodologia que permita a implementação deste modelo de negócio. Como tal, será possível a criação de um modelo de classificação do cabaz alimentar dos clientes através dos dados recolhidos por um cartão de fidelização, capaz de produzir efeitos benéficos – isto é, descontos – no prémio de um seguro de saúde?

3. Metodologia

O objetivo do presente capítulo é apresentar a metodologia utilizada para responder aos objetivos de investigação definidos – i) *contribuir para a definição de uma metodologia de classificação dos bens alimentares transacionados*; e ii) *analisar se a metodologia desenvolvida no primeiro ponto é capaz de ser replicável para o universo das transações e clientes fidelizados*.

Primeiramente, será feita uma apresentação da metodologia de investigação adotada na concretização do projeto. De seguida, serão identificadas as metodologias aplicadas no desenvolvimento do projeto, sendo que a primeira se debruça sobre o estudo dos hábitos de consumo associados a um cartão de fidelização, no sentido de desenvolver um modelo de classificação do cabaz de consumo alimentar existente no retalhista. Por último, será estudada a aplicação do modelo definido numa amostra aleatória, para comprovar a sua veracidade e utilização, bem como propor uma medida quantitativa que permita a implementação de descontos associados ao prémio de um seguro. Importa salientar que um dos propósitos deste trabalho se prende com a modelização, e assim criar condições para que o trabalho possa ser replicado de forma sistemática em análises de futuras transações realizadas pelos clientes.

3.1. Metodologia de Investigação

Tendo em conta o cariz inovador da investigação, foi conduzida uma análise exploratória, que procura descobrir o contexto atual em que o problema se insere e obter conhecimento sobre determinado tópico, sendo assim flexível e ajustável (Saunders et al., 2011). Existem diversas formas de conduzir uma investigação

exploratória, desde revisão da literatura existente, à realização de entrevistas com profissionais relacionados com o tema/área ou realização de estudos de mercado. Deste modo, uma análise exploratória recorre, muitas das vezes, à observação para recolher dados primários, o que envolve a visualização sistemática, registo, descrição, análise e interpretação do comportamento das pessoas (Saunders et al., 2011)

Tendo presente a questão de investigação proposta, bem como os objetivos de investigação definidos, este projeto incluiu a realização da análise documental existente em plataformas científicas académicas – como a *Web of Science*, Repositório *Veritati* e *Google Scholar*; bem como em plataformas da empresa responsável pelo estudo. Por fim, foram realizadas entrevistas informativas com diversos colegas do grupo onde o estudo decorreu, no sentido de: i) aferir a existência de informação estruturada sobre as características nutricionais dos produtos; ii) conhecer a estrutura da base de dados do retalhista; iii) obter uma amostra representativa aleatória com informação transacional de clientes fidelizados.

Em segundo lugar, esta análise, de natureza qualitativa, subdivide-se em duas metodologias: uma primeira, onde, ainda que baseada em critérios numéricos, é proposta uma classificação do cabaz de consumo alimentar do retalhista; e uma segunda metodologia que pretende comprovar a aplicabilidade do modelo desenvolvido e estudar as relações entre as variáveis através de gráficos e análises (Saunders et al., 2011). A escolha por uma abordagem qualitativa está comumente associada a uma filosofia interpretativa, visto que implica a compreensão dos resultados obtidos para poder interpretar os fenómenos estudados. Deste modo, a pesquisa qualitativa recorre a técnicas de recolha e análise de dados e a criação de conhecimento é conseguida através do estudo das relações entre as variáveis (Saunders et al., 2011).

3.2. Conceção do Modelo-SN

Para o desenvolvimento e estruturação do modelo de classificação do cabaz de consumo alimentar existente no retalhista – que intitulei como Modelo-SN, Modelo do Semáforo Nutricional – importa salientar que o retalhista já utiliza, há aproximadamente 10 anos, um sistema de classificação através do código de cores de grande parte dos seus produtos pré-embalados de marca própria. Contudo, uma vez que o retalhista não comercializa unicamente produtos marca própria, para poder classificar todo o cabaz de consumo alimentar seria necessária a replicação do modelo para os produtos das marcas parceiras. Adicionalmente, o sistema utilizado pelo retalhista não se encontra estruturado nas suas bases de dados.

Como tal, foi realizado um levantamento das Direções Comerciais existentes nas lojas do retalhista para as categorias de vens alimentares, com vendas ao longo do ano 2018. Entende-se como Direção Comercial o nível ou segmento de maior agregação existente para o produto. De seguida, a análise foi restrita aos departamentos alimentares. É apresentado na **Figura 2**, um esquema que permite compreender melhor agregação de níveis de classificação de produtos do retalhista, desde o departamento comercial até ao código do produto em si. Na **Tabela 7** é apresentado um exemplo da respetiva estrutura existente em base de dados, para o caso concreto da Sub Categoria *Batatas Fritas*.

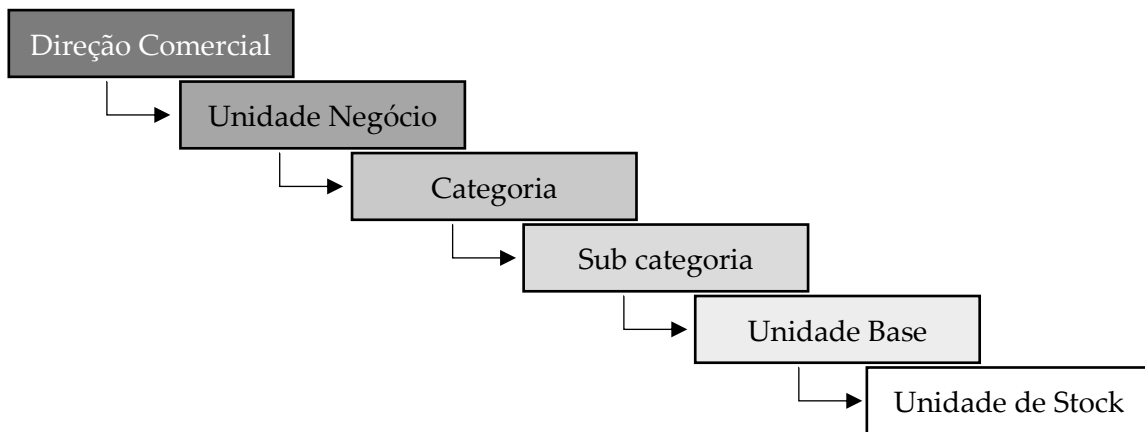


Figura 2 – Níveis de agregação de produtos existentes no Retalhista

Direção comercial	Unidade de negócio	Categoria	Sub Categoria	Unidade Base
10 - Alimentar	01 - Merceria Salgada	0101 - Aperitivos	010101 - Batata Frita	01010101 - Sal Lisa
10 - Alimentar	01 - Merceria Salgada	0101 - Aperitivos	010101 - Batata Frita	01010102 - Sal Ondulada
10 - Alimentar	01 - Merceria Salgada	0101 - Aperitivos	010101 - Batata Frita	01010103 - Sabores Lisa
10 - Alimentar	01 - Merceria Salgada	0101 - Aperitivos	010101 - Batata Frita	01010104 - Sabores Ondulada
10 - Alimentar	01 - Merceria Salgada	0101 - Aperitivos	010101 - Batata Frita	01010105 - Artesanais Palha
10 - Alimentar	01 - Merceria Salgada	0101 - Aperitivos	010101 - Batata Frita	01010106 - Artesanais Lisa
10 - Alimentar	01 - Merceria Salgada	0101 - Aperitivos	010101 - Batata Frita	01010108 - Bf Take-Away
10 - Alimentar	01 - Merceria Salgada	0101 - Aperitivos	010101 - Batata Frita	01010109 - Light/Saudáveis
10 - Alimentar	01 - Merceria Salgada	0101 - Aperitivos	010101 - Batata Frita	01010110 - Gourmet
10 - Alimentar	01 - Merceria Salgada	0101 - Aperitivos	010101 - Batata Frita	01010111 - Bat. Fri. Outras

Tabela 7 – Exemplo da estrutura da base de dados do retalhista

Concluída a primeira fase, numa primeira análise foram encontrados aproximadamente 1 900 000 0000 de Unidade de Stock alocados a cerca de 2 000 segmentos de Unidade Base. Deste modo, definiu-se que para alcançar o primeiro objetivo de investigação, isto é, para definir o Modelo-SN, a análise será realizada

ao nível de agregação Unidade Base (ou níveis de agregação superiores) e não à Unidade de Stock. Adicionalmente, a realização da análise a um nível de agregação superior permite a replicação do modelo de uma forma mais eficiente, uma vez que a introdução de novos produtos em categorias existentes está contemplada no modelo.

Por conseguinte, reconheceu-se a importância da definição das seguintes fases: análise e validação da abrangência dos segmentos; e categorização e estruturação dos mesmos.

A análise e validação da abrangência dos segmentos requereu uma pesquisa mais detalhada às Unidades de Negócio, onde foram eliminadas do estudo categorias de produtos não alimentares. São exemplo de categorias desconsideradas os suplementos, guardanapos e lenços de papel, produtos de higiene e de limpeza. Para validar a abrangência das categorias existentes em base de dados, foi realizada uma análise prévia dos produtos incluídos nos segmentos com objetivo de aferir se, pela identificação apenas das Unidades Base era possível definir um modelo de estudo e, paralelamente, avaliar uma categorização por este nível ou superior. A Tabela 7, apresenta a análise realizada à Sub Categoria *Batatas Fritas*, onde é possível identificar, através da nomenclatura existente para a Unidade Base, qual a informação nutricional a analisar, e daí reconhecer a possibilidade de agregação das Unidades Base existentes, dado que inclui produtos com informação nutricional muito semelhantes.

A terceira fase, categorização e estruturação dos segmentos alimentares existentes no retalhista, foi realizada através do cruzamento entre, por um lado, diversas pesquisas à plataforma digital do retalhista, onde são disponibilizados ao consumidor, para um número considerável de produtos, a informação nutricional dos mesmos; e, por outro lado, o estudo de Dunford et al. (2017), que

apresenta um modelo para classificação dos produtos pré-embalados de acordo com a quantidade nutricional existente nos mesmos.

Deste modo, o Modelo-SN é analisado para todas as Unidades Base alimentares do retalhista da seguinte forma:

- i. Pesquisa do produto relativo à Unidade Base na plataforma de compras online do retalhista;
- ii. Verificação da existência de informação nutricional sobre o mesmo;
- iii. Comparação das quantidades nutricionais para cada um dos quatro nutrientes propostos com o modelo de Dunford et al. (2017);
- iv. Atribuição da pontuação por nutriente, conforme proposto por Dunford et al. (2017);
- v. Atribuição da classificação global do produto.

É importante mencionar que o modelo de Dunford et al. (2017) distingue a classificação de produtos consoante sejam alimentos ou bebidas, pelo que a presente metodologia replicou a proposta. Adicionalmente, enquanto os autores classificam o limite considerado para uma alimentação saudável (até 6 pontos num total de 12), este limite não é proposto para a distinção entre uma alimentação intermédia ou desfavorável. Deste modo, consideraram-se classificações superiores a 8 pontos como desfavoráveis. É apresentado na **Tabela 8** um exemplo da classificação feita para a base de dados estudada, onde se identifica a informação nutricional para um produto pertencente a cada Unidade Base.

No entanto, uma vez que a análise de Dunford et al. (2017) recorre a um modelo de classificação de produtos pré-embalados, foi necessário recorrer a outras investigações para classificar produtos frescos, como frutas, legumes, carne, peixe, entre outros. Trichopoulou, Bamia, & Trichopoulos (2009) caracterizam a dieta Mediterrânica como uma dieta que prima pelo elevado consumo de legumes, fruta, frutos secos, peixe, marisco e cereais, e por um

reduzido consumo de carnes e derivados e gorduras. Assim, a classificação dos frutos e legumes com uma categoria verde parte da verificação de que estes alimentos são benéficos para a saúde e cruciais para a adoção de uma alimentação saudável (Arechavala et al., 2016; Conner et al., 2017; Holubcikova et al., 2016; Trichopoulou et al., 2009; Voráčová et al., 2016). Adicionalmente, foi definida uma classificação com a categoria verde para os alimentos peixe e marisco (Trichopoulou et al., 2009).

Unidade base	SKU exemplo	Lípidos		Saturados		Açúcares		Sal		Total
		por 100 g	Score	por 100 g	Score	por 100 g	Score	por 100 g	Score	
01010101 - Sal lisa	2005831	34 g	R	14 g	R	0,5 g	G	1,2 g	R	R
01010102 - Sal ondulada	2005830	34 g	R	14 g	R	0,7 g	G	1,2 g	R	R
01010103 - Sabores lisa	6328753	31 g	R	12 g	R	2,5 g	G	1,3 g	R	R
01010104 - Sabores ondulada	6328736	32 g	R	11 g	R	0,6 g	G	1,5 g	R	R
01010105 - Artesanais palha	4129602	35 g	R	16 g	R	0,8 g	G	1,0 g	R	R
01010106 - Artesanais lisa	2060425	32 g	R	13 g	R	0,5 g	G	1,5 g	R	R
01010108 - bf take-away	5359323	33 g	R	4 g	Y	0,5 g	G	1,3 g	R	R
01010109 - light/saudaveis	3699358	24 g	R	12 g	R	0,5 g	G	1,2 g	R	R
01010110 - gourmet	2092498	32 g	R	3,3 g	Y	0,6 g	G	1,5 g	R	R

Tabela 8 – Exemplo ilustrativo de classificação do Modelo-SN

Em que:

G – Classificação verde, equivalente a quantidades reduzidas do nutriente ou pontuação total;

Y – Classificação amarela, equivalente a quantidades intermédias do nutriente ou pontuação total;

R – Classificação vermelha, equivalente a quantidades elevadas do nutriente ou pontuação total.

Para classificar alimentos na Unidade de Negócio *Talho*, é importante distinguir entre os grupos de carne, pois existe uma grande variação nutricional entre a carne vermelha e carne branca. Assim, de acordo com Richi et al. (2015), o consumo de carne vermelha e de carne processada está associado a um aumento do risco de mortalidade, surgimento de doenças cardiovasculares, cancro colorretal e diabetes tipo II, tanto em homens quanto em mulheres. Por outro lado, esta associação não se identifica na carne branca. Com base na informação apresentada por Richi et al. (2015) e Trichopoulou et al. (2009), classificaram-se os alimentos de carne branca com a categoria amarela e os alimentos de carne vermelha ou processada com a categoria vermelha.

Por último, ao analisar as bebidas alcoólicas, como vinhos e espirituosas, tendo por base o Modelo-SN, a classificação atribuída seria verde, pois o único nutriente considerado na maioria destes produtos como vermelho era o nível de açúcares. Ora, de acordo com a Organização Mundial de Saúde, o consumo de bebidas alcoólicas é considerado prejudicial para a saúde e consumido em quantidades excessivas pode causar doenças no fígado, cancro, doenças cardíacas e mentais⁴. Como tal, os alimentos que se encontram nestes grupos foram classificados com a categoria vermelha.

É ainda importante mencionar que para o presente estudo se optou por classificar determinadas categorias com uma pontuação nula. Tal classificação está relacionada com a amplitude de produtos de classificações nutricionais díspares nas Sub Categorias associadas, ou com o facto de se tratarem de refeições pré-preparados e, portanto, a sua informação nutricional não é fornecida. As Sub categorias classificadas com pontuação nula são: i) *180103 – Refeições*, onde estão incluídos produtos como sopas e saladas, pizzas, sandes, refeições pré-embaladas, acompanhamentos, entre outros disponibilizados aos clientes através de balcões de atendimento; ii) *180104 – Parcerias*, onde constam

⁴ https://www.who.int/topics/alcohol_drinking/en/ (acedido a 19.01.2019)

refeições, essencialmente de cozinha japonesa e chinesa; e iii) 180201 – Refeições – que inclui produtos semelhantes aos produtos do ponto i), mas estão previamente embalados e disponibilizados aos clientes em balcões de livre serviço.

A conjugação de todas as classificações feitas compõe, deste modo, o Modelo-SN, e permite alcançar o primeiro objetivo de investigação.

3.3. Teste do Modelo-SN

Com o intuito de encontrar contribuições futuras para o desenvolvimento de um seguro orientado pelo *data-driven*, a segunda metodologia utilizada pretende aplicar o modelo definido no subcapítulo anterior numa amostra de dados transacionais do retalhista. Para tal, foi selecionada uma amostra aleatória das transações realizadas com cartão de fidelização de uma fração da base de clientes do retalhista, de dezembro de 2017 a novembro de 2018. Para complementar a análise, foram extraídas tabelas complementares, com dados de dimensão sobre os clientes. Deste modo, foram selecionadas três tabelas, com as seguintes particularidades:

- Informação_Transacional: onde constam, para cada cliente, todas as transações fidelizadas realizadas durante o período de análise. Adicionalmente, o ficheiro contém informação relativa à data e local da transação, código da Unidade de Stock, fator de conversão do produto (isto é, indica o número de quantidades vendidas do produto na transação ou peso do mesmo), preço do produto e desconto associado;
- Dimensão_Produto: onde é indicado, para cada código de Unidade de Stock, a desagregação do produto até ao departamento ou Direção Comercial associada. É através deste ficheiro que é possível aplicar o

Modelo-SN, classificado ao nível da Unidade Base, para todas as transações da amostra;

- Dimensão_Cliente: que identifica, para cada cliente segmentado no ficheiro Informação_Transaccional, quais são as variáveis socioeconómicas associadas. Para a presente análise, foram extraídos dados relativos ao género, ano de nascimento, número de membros no agregado familiar e identificador do segmento estilo de vida (medida calculada através de desenvolvimentos *advanced analytics*).

Dada a dimensão dos ficheiros em estudo, a análise do Modelo-SN foi desenvolvida com recurso a *Power BI Desktop*. O *Power BI Desktop* é uma aplicação do *Office 365* para visualização de dados complexos e que permite a importação, edição e modelação de dados, bem como a criação de *reports*, ecrãs de visualização, e a partilha destes resultados na *cloud*, por forma a que possam ser consultados quer em computador como *tablets* e *smartphones*.

Numa fase prévia à importação das tabelas em *Power BI Desktop*, foi adicionado ao ficheiro Dimensão_Produto, o Modelo-SN. Esta associação, entre o *SKU* de cada produto seleccionado e a classificação do modelo ao nível da Unidade Base foi feita através do *Excel*. Adicionalmente, ao analisar os dados extraídos no ficheiro Dimensão_Cliente, foram identificados clientes onde o campo relativo ao número de membros do agregado familiar era questionável ou estava a branco. Deste modo, para uniformizar a análise, os clientes segmentados com campo *Family_Members* a branco foram classificados como tendo um agregado familiar de 1 pessoa, e os clientes segmentados com valores iguais ou superiores a 5, foram classificados como tendo um agregado familiar de 5 pessoas.

Findas as uniformizações prévias, foram importados os ficheiros em *Power BI Desktop*, bem como estabelecidas as variáveis-chave que permitem o relacionamento entre as tabelas. Especificamente, foi definida a relação de um para muitos entre a tabela Dimensão_Cliente e a tabela de Informação_Transaccional através da variável comum em ambas – *Customer_Account_Nr_Mask*; e a relação de um para muitos entre a tabela Dimensão_Produto e a tabela de Informação_Transaccional através da variável presente em ambas – *SKU*. A **Figura 3** apresenta um excerto retirado da aplicação *Power BI Desktop*, onde as mesmas relações foram definidas.

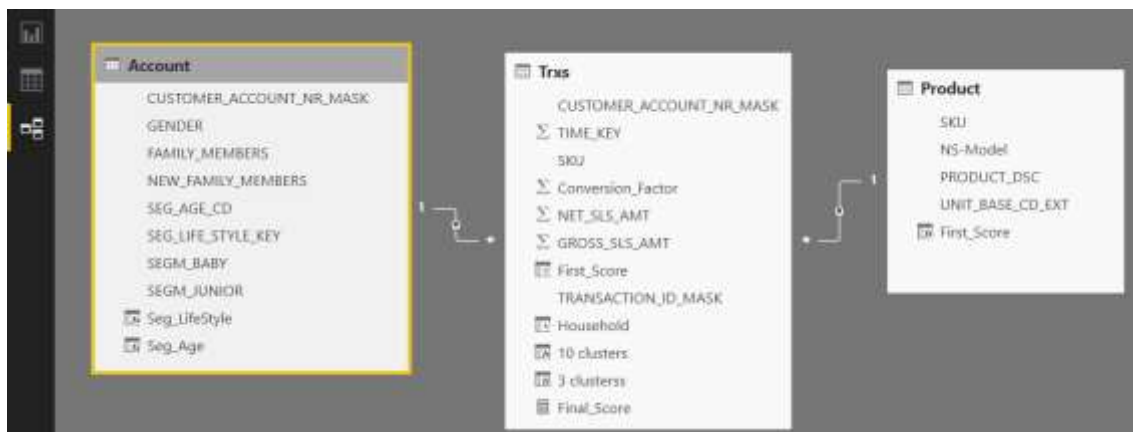


Figura 3 - Relações estabelecidas nas tabelas de análise

A terceira fase da análise do modelo aplicado à amostra aleatória implicou a criação de colunas complementares de apoio ao estudo, bem como a definição de métricas adicionais. Quanto à criação de colunas complementares, foram definidas as seguintes colunas:

- Na tabela Produto, foi criada uma medida de *score* associada ao Modelo-SN. Uma vez que este modelo atribui uma classificação qualitativa para cada produto, é necessário associar uma medida quantitativa para cada categoria, para que seja possível observar as disparidades de consumo em *Power BI*. Como tal, a classificação das métricas baseou-se na pontuação definida por Dunford et al. (2017) para classificação de cada nutriente: produtos com categoria verde

recebem 1 ponto; amarela recebem 2 pontos; vermelha recebem 3 pontos. Por fim, produtos alimentares excluídos opcionalmente recebem zero pontos.

- Na tabela Transações, foi criada uma coluna que, para o *First_Score* criado na tabela de Produto, associe todas as transações importadas.
- Na tabela Transações, foi criada uma coluna para associar o número de membros do agregado familiar à transação. Deste modo, será possível considerar esta variável na classificação do cabaz de consumo alimentar. Especificamente, considere-se a comparação entre 1 kg de carne consumida por uma família de 2 pessoas e a mesma quantidade consumida por uma família de 5 pessoas.

Por fim, foi criada na tabela de Transações uma métrica de *Score Final* do cliente cujo objetivo é permitir a classificação do cabaz alimentar por cliente, e que considera todas as transações alimentares fidelizadas, aplicando o Modelo-SN e ajustando o resultado pelo fator de conversão de cada produto e número de membros do agregado familiar. A métrica definida apresenta a seguinte equação:

$$Final_{score} = \frac{\sum(Si \times FC \div H)}{\sum Trx}$$

Em que:

Si – *Score* inicial;

FC – Fator de Conversão;

H – Número de membros do agregado familiar;

Trx – Transações de categoria alimentar consideradas no Modelo-SN.

Concluída a definição da métrica de cálculo da pontuação final a atribuir a cada cliente pelo total do seu cabaz alimentar, tornou-se possível analisar a distribuição dos Scores Finais na amostra. O estudo da distribuição dos perfis de cliente, consoante apresentem padrões de consumo alimentares mais ou menos saudáveis, permite a definição de segmentos de cliente que podem futuramente

ser aplicados na atribuição de um desconto associado ao prémio de um seguro de saúde ou de vida. Contudo, e conforme apresentado em capítulos anteriores, a literatura existente não apresenta nenhuma proposta para a aplicação deste modelo, pelo que as elações apresentadas pretendem contribuir para estudos futuros, deixando bem claro que ainda não estão comprovadas.

Estando a análise a ser realizada em *Power BI Desktop*, a aplicação possui uma funcionalidade, que requer a criação prévia de uma tabela que considere todos os clientes selecionados e o *Score Final*, e que permite a definição automática de clusters, sendo que o utilizador necessita apenas de definir um intervalo de segmentos, no máximo 15, que pretende obter. Para esta segmentação foram ordenados os clientes por *Score Final* e, posteriormente, selecionados 10 clusters.

Finda a modelação dos dados, tornou-se então possível analisar os resultados do modelo de estudo desenvolvido. A análise dos resultados será apresentada no capítulo seguinte.

3.4. Sumário

O desenvolvimento de um seguro cujo prémio pago pelos consumidores possa ser beneficiado de acordo com o seu perfil – neste caso, com o perfil de consumo alimentar – requer a definição de metodologias de implementação robustas, que não estão, até à data, desenvolvidas em qualquer seguro praticado no mercado. Deste modo, o presente projeto propôs-se a estudar e desenvolver um modelo de classificação do cabaz de consumo alimentar associado a um cartão de fidelização abrangente, no sentido de iniciar os desenvolvimentos associados a este modelo de seguro.

De forma a estudar este tema, foi, primeiramente, feito um levantamento das categorias de consumo existentes no retalhista, e numa segunda fase, atribuído

uma medida de classificação de cada categoria e respectivos produtos incluídos, tendo por base investigações científicas publicadas. Terminada a estruturação do modelo, foi possível aplicar o mesmo a uma amostra aleatória durante um ano de transações fidelizadas. Dada a dimensão dos dados a analisar, foi utilizada a aplicação *Power BI Desktop* para observar o comportamento do Modelo-SN aplicado na amostra.

4. Apresentação e discussão dos resultados

Este capítulo tem como objetivo apresentar os resultados identificados com a definição e aplicação do Modelo-SN pretendendo, primeiramente, comparar os resultados obtidos com os hábitos de consumo alimentares estudados pelos autores apresentados na Revisão Bibliográfica e, numa segunda fase, comparar os *clusters* segmentados pela aplicação, no sentido de propor uma medida de classificação do cabaz, capaz de ser aplicada num desconto associado a um seguro.

4.1. Apresentação dos resultados

Para definir o Modelo-SN, foi necessária a classificação de aproximadamente 1 500 Unidades Base. Contudo, conforme referido anteriormente, em determinados segmentos de produto, esta classificação foi realizada a um nível de agregação superior por ser relativa a categorias de alimentos com características nutricionais muito semelhantes. O Apêndice I apresenta a classificação proposta para o Modelo-SN.

Finda a classificação realizada ao nível da Unidade Base, obteve-se um total de 44% de classificações na categoria verde, 16% na categoria amarela, e 39% na categoria vermelha. As categorias com classificação opcional nula representam 1% da base.

No que respeita à análise da amostra aleatória de transações fidelizadas com cartão de fidelização, foram classificadas de acordo com o Modelo-SN, e por pertencerem a categorias alimentares, cerca de 77% do total de transações

extraídas. Paralelamente, no ficheiro relativo aos dados Dimensão_Produto, foram consideradas para a classificação do modelo por pertencerem a categorias alimentares estudadas, 21% dos códigos de produto totais.

Considerando que a amostra selecionada representa uma porção da base de clientes com cartão de fidelização, foram extraídos para análise cerca de 74 000 clientes. Destes, 63,3% são elementos do género feminino e 36,7% são elementos do género masculino. Quanto ao perfil dos clientes segmentados, identificou-se a seguinte distribuição quanto ao número de membros do agregado familiar: 26,6% dos clientes são os únicos membros do seu agregado familiar; 24,1% têm 2 membros; 21,8% têm 3 membros; 18,3% têm 4 membros; e 9,2% têm 5 membros. Por fim, analisando o perfil dos clientes quanto à sua segmentação de estilo de vida, obteve-se uma distribuição de 19,5% para os clientes identificados no segmento SLS_1; 16,8% para o SLS_2; 10,1% para o SLS_3; 10,3% para o SLS_4; 13,9% para o SLS_5; 12,9% para o SLS_6; 14,0% para o SLS_7; e 2,4% o SLS_8.

Ao aplicar o *Score* Inicial no total de transações realizadas – isto é, dados contidos na tabela de Informação_Transacional – verificou-se que 50,7% das transações realizadas obtiveram uma classificação de verde; 18,5% uma classificação amarela; e 30,4% uma classificação vermelha. O peso das categorias alimentares excluídas do modelo no total das transações é inferior a 0,4%. Por outro lado, ao analisar os produtos incluídos na amostra segmentada, numa ótica de contabilização única – ou seja, códigos irrepetíveis incluídos na tabela Dimensão Produto – verificou-se um peso de 36,8% dos produtos categorizados a verde; 15,5% a amarelo; 45,1% a vermelho; e 2,6% a nulo. A Tabela 9 apresenta os resultados, em percentagem, obtidos na amostra por unidade de *stock* distinto e no total de transação, classificado de acordo com o Modelo-SN.

Modelo-SN	Total de transações (SKUs)	SKUs distintos
Verde – G	50,7%	36,8%
Amarelo – Y	18,5%	15,5%
Vermelho – R	30,4%	45,1%
0	0,4%	2,6%
Total	100%	100%

Tabela 9 – Distribuição do Modelo-SN por transações e SKU

Considerando a aplicação da métrica *Final_Score* no total da população segmentada, e discriminando o resultado de acordo com a métrica *Final_Score* por cada uma das três categorias propostas pelo Modelo-SN obteve-se um score médio de 0,83 para produtos classificados a verde; 1,40 para produtos classificados a amarelo; e 2,23 para produtos classificados a vermelho (**Figura 4**).

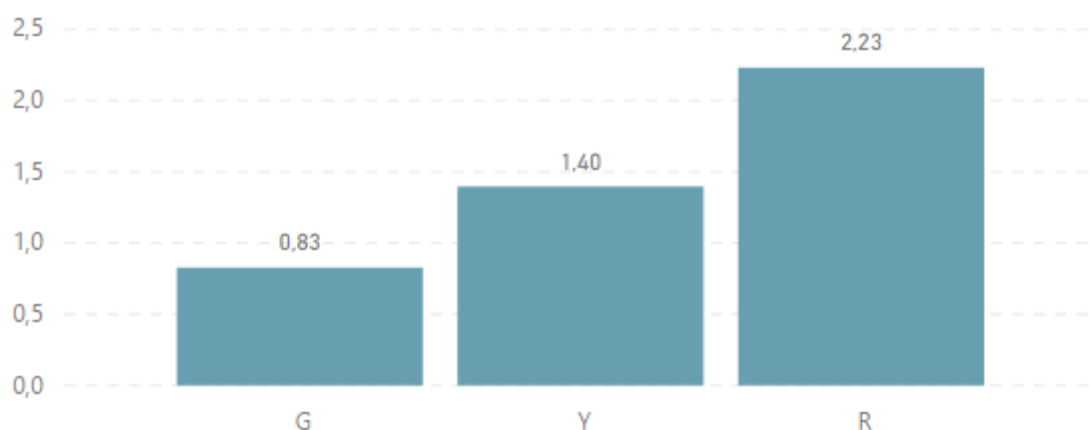


Figura 4 - Média do Score Final discriminado de acordo com o Modelo-SN

Analisando os resultados de acordo com o perfil de clientes, foi obtido um *Score Final* médio de 1,25 para elementos do género feminino; e de 1,50 para elementos do género masculino (**Figura 5**).

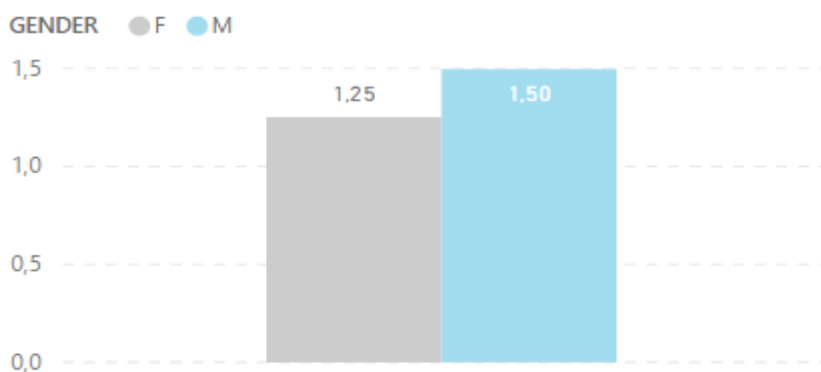


Figura 5 - Média do Score Final discriminado por género

Em que:

F – Elementos do género feminino

M – Elementos do género masculino

Comparando o *Score Final* com a medida de segmentação do estilo de vida proposta pelo Retalhista, obtiveram-se os seguintes resultados: uma pontuação de 1,79 de para o segmento SLS_1; 1,42 para o segmento SLS_2; 1,40 para o segmento SLS_3; 1,34 para o segmento SLS_4; 1,33 para o segmento SLS_5; 1,29 para o segmento SLS_6; 1,31 para o segmento SLS_7; e 1,57 para o segmento SLS_8 (**Figura 6**).

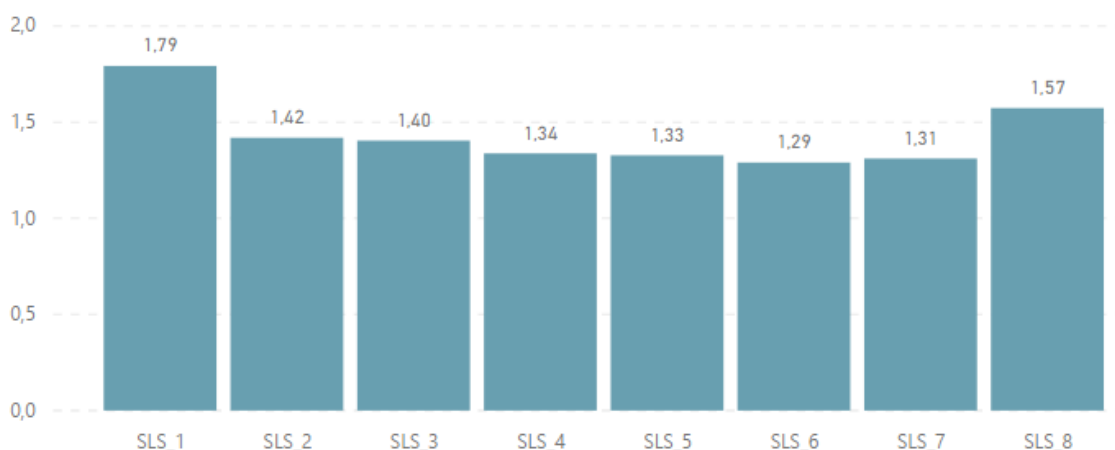


Figura 6 - Média do Score Final discriminado pelo Segmento Estilo de Vida

Por fim, a definição de clusters em *Power BI Desktop* permitiu, numa primeira fase, identificar o *Score Final* médio por *cluster*, e intervalo de resultados mínimo e máximo que cada *cluster* inclui. Numa segunda fase, e devido à exportação de cada grupo de cluster para um documento *Excel*, foi possível identificar o número de clientes segmentados para cada *cluster*, reconhecendo assim o peso de cada grupo no total. Na **Figura 7** são apresentados os 10 clusters definidos de acordo com a aplicação, ordenados, de forma decrescente, acordo com o *Score Final* médio do segmento. Na **Tabela 10** é apresentado o peso de cada cluster no total da amostra.

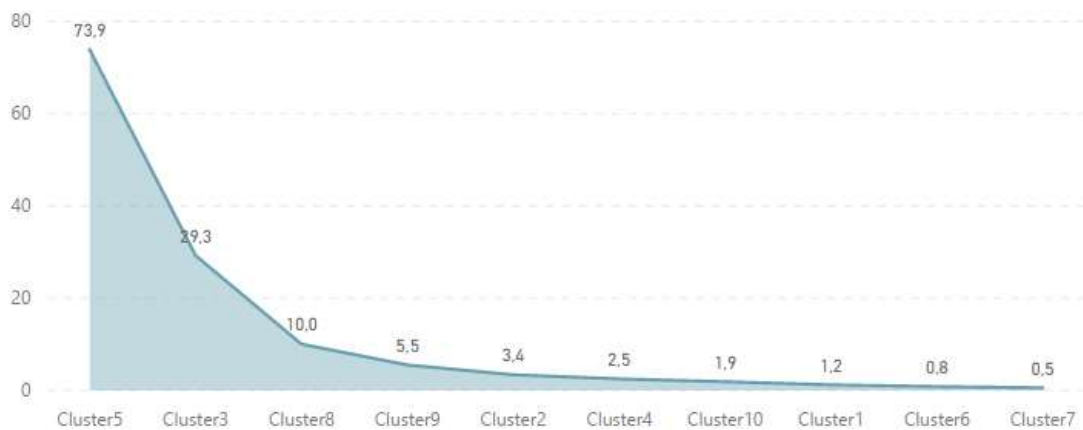


Figura 7 – Distribuição dos clusters de acordo com o total de clientes e Score Final

Cluster	Score Mínimo	Score Máximo	Nº Clientes	Peso no total da base
C7	-	0,66	15 742	20,42%
C6	0,66	1,00	18 042	23,40%
C1	1,01	1,50	18 598	24,12%
C10	1,51	2,15	9 233	11,97%
C4	2,16	2,94	10 157	13,17%
C2	2,95	4,51	4 433	5,75%
C9	4,52	8,46	722	0,94%
C8	8,50	18,79	140	0,18%
C3	19,37	42,00	34	0,04%
C5	56,17	75,15	4	0,01%

Tabela 10 - Distribuição dos clientes de acordo com os clusters definidos

4.2. Discussão

Quando pensamos na diversidade de produtos alimentares disponíveis aos consumidores em qualquer retalhista, é possível identificar uma amplitude extensa de categorias de artigos, em constante renovação. Sendo assim, a determinação do Modelo-SN através da análise realizada às categorias alimentares do retalhista, permite deprender duas ideias correlacionadas: a primeira evidencia a multiplicidade de produtos presentes nas categorias; a segunda, realça a potência da análise feita às categorias, visto que permite a continuidade do modelo sempre que são adicionados novos artigos, sem que se crie uma nova categoria.

Em primeiro lugar, a construção do Modelo-SN permitiu a segmentação das categorias alimentares em 3 padrões de consumo – verde, para alimentos mais

saudáveis; amarelo, para alimentos intermédios; e vermelho, para alimentos pouco saudáveis. Enquanto no estudo realizado por Dunford et al. (2017) aos produtos comercializados pelos retalhistas americanos, 30,1% dos produtos analisados obtiveram uma classificação saudável. Em contraste, este estudo obteve um peso de 44% para os produtos classificados com a categoria verde nos produtos comercializados por um dos maiores retalhistas portugueses. Deste modo, numa primeira análise poderá inferir-se que os retalhistas alimentares portugueses tendem a oferecer produtos mais saudáveis. Contudo, enquanto os primeiros autores restringem a sua análise a produtos pré-embalados, tendencialmente menos saudáveis; o segundo propõe uma classificação para todo o cabaz alimentar, incluindo produtos frescos e geralmente mais saudáveis.

Em segundo lugar, ao comparar a proporção de produtos classificados a verde na amostra aleatória com o peso desses produtos no total das transações segmentadas, é possível reparar que enquanto os produtos com categorização saudável representam 37% do total de produtos distintos, estes são consumidos com uma maior frequência e repetição, apresentando um peso de 51% no total das transações. Por oposição, os consumidores tendem a consumir uma quantidade menor, mas mais variável de produtos com categoria vermelha – sendo que representam um peso no cabaz de consumo alimentar dos clientes de 19%, mas um peso nas categorias de consumo segmentadas de 45%. Uma explicação possível para estes resultados prende-se com o facto de os consumidores não apresentarem significativas variações nos produtos que compram de forma rotineira – como frutas, legumes e leite – mas tendem a seleccionar opções diversificadas no que respeita a produtos indulgentes.

Em terceiro lugar, aplicação do *Score Final* na amostra aleatória permitiu aferir que os elementos do género feminino apresentam uma alimentação mais favorável que os elementos do género masculino. Ora, este resultado aparenta

estar de acordo com os resultados obtidos por Arechavala et al. (2016), Costa-Tutusaus & Guerra-Balic (2016) e Holubcikova et al. (2016), conforme referido no capítulo que estuda os hábitos alimentares e o género. Importa, contudo, relembrar que, uma vez que o presente estudo é feito através das compras fidelizadas, não é possível afirmar com certeza que a transação final tenha sido realizada pelo detentor do cartão, ou por outro membro do agregado familiar, de outro género.

Ainda sobre a aplicação do Score Final, mas desta vez de acordo com a classificação de *advanced analytics* proposta pelo retalhista – segmento estilo de vida – identificaram-se 2 segmentos com um Score Final mais saudável, os SLS_6 e os SLS_7. Sendo assim, apesar de a literatura apresentada analisar a relação entre os hábitos alimentares e o status socioeconómico ou escala de afluência familiar, é possível retirar algumas interpretações destes resultados. Arechavala et al. (2016) e Voráčová et al. (2016) defendem que as famílias mais afluentes tendem a ser mais preocupadas com a alimentação. Paralelamente, o facto de os consumidores estarem conscientes da nutrição dos alimentos influencia o consumo dos mesmos (Holubcikova et al., 2016; Voráčová et al., 2016). Deste modo, a pontuação final média atribuída aos dois segmentos poderá estar assente nestas premissas. Por outro lado, o preço dos alimentos impacta o consumo dos mesmos (Arechavala et al., 2016; Costa-Tutusaus & Guerra-Balic, 2016; Voráčová et al., 2016).

Por fim, a determinação de *clusters* pelo *Power BI Desktop* permitiu a formulação de uma proposta de classificação do cabaz de consumo alimentar, capaz de permitir a aplicação de descontos associados a um seguro de saúde ou vida. Ainda que não faça parte do âmbito do trabalho – a discussão sobre o modelo de *pricing* aplicado – os parâmetros foram definidos no sentido de permitirem a customização do prémio, bem como a aplicação de descontos maiores consoante o Score Final seja mais reduzido. Deste modo, a **Tabela 11**

apresenta o Modelo de Classificação Global, e pressupõe a aplicação de descontos associados para um *Score Final* inferior ou igual a 1.50.

Clusters	Score Final Mínimo	Score Final Máximo	% de Clientes na amostra	Segmentação Proposta
C7	-	0.70	20.50%	A
C6	0.71	1.00	23.50%	B
C1	1.01	1.50	24.00%	C
C4, C10	1.51	3.00	25.00%	D
C2, C3, C5, C8, C9	3.01	ilimitado	7.00%	E

Tabela 11 - Segmentação do Cabaz de consumo cabaz de implicar descontos associados ao prémio

Tendo em conta o cariz inovador do projeto, bem como lacuna associada à bibliografia existente, considera-se relevante referir que a construção do Modelo de Classificação Global teve como base o peso dos clientes na amostra, como forma de definir os limites mínimos e máximos associados ao *Score Final* de cada cliente.

5. Conclusão

Desde uma fase inicial da realização deste projeto que foi possível identificar uma lacuna de estudos científicos sobre esta matéria – correlação entre padrões de consumo e prémio de seguros. Por um lado, tornou o trabalho mais estimulante, pois em muitas alturas foi necessário apontar caminhos e soluções para áreas do conhecimento que não estavam devidamente estudadas. Por outro lado, relembra a necessidade de o Modelo-SN se encontrar apto a trabalhar com múltiplas formas de classificar os alimentos, e que isso possa ser feito de forma dinâmica. O presente capítulo apresenta as principais implicações do desenvolvimento do estudo, assim como as forças e limitações identificadas e, por fim, as perspetivas de trabalhos futuros.

5.1. Implicações

Ao longo deste estudo, aprofundou-se o conhecimento associado à correlação existente entre os padrões de consumo alimentares e o prémio de um seguro. Concluído o projeto, e tendo em conta o objetivo de investigação definido no início do trabalho – *contribuir para o desenvolvimento de um novo modelo de negócio de seguros, cuja determinação do risco associado considere os hábitos alimentares dos clientes* – é possível afirmar que as análises realizadas possibilitaram o alcance do mesmo.

Avaliando o primeiro objetivo complementar, i) *contribuir para a definição de uma metodologia de classificação dos bens alimentares transacionados*, poder-se-á dizer que o presente trabalho permitiu o desenvolvimento de uma primeira investigação que propõe a formulação de um modelo de classificação dos hábitos

de consumo alimentares associados a um cartão de fidelização. A formulação do modelo permite, assim, destacar a enormidade informação para estruturar. No entanto, pode considerar-se que a modelação foi conseguida, concluindo num modelo simples e intuitivo: onde produtos mais saudáveis são classificados a verde, produtos intermédios a amarelo, e os produtos menos saudáveis são classificados a vermelho. Por último, evidencia-se a potencialidade de o trabalho poder ser enriquecido se, ao Modelo-SN, se acrescentarem outras componentes, apresentadas no subcapítulo dedicado às perspetivas futuras.

No que diz respeito ao segundo objetivo complementar, ii) *analisar se a metodologia desenvolvida no primeiro ponto é capaz de ser replicável para o universo das transações e clientes fidelizados*, o desenvolvimento do projeto incluiu a definição de uma métrica quantitativa com a finalidade de permitir a classificação do cabaz alimentar por cliente, bem como a construção de *clusters* que permitem identificar segmentos de aplicação de descontos. Como tal, a métrica *Score Final* permitiu a identificação de grupos de clientes com maior propensão em aderir a um seguro como este – os SLS_6 e os SLS_7 – que, por adotarem dietas mais saudáveis, se acredita que poderão reconhecer mais facilmente os benefícios associados ao seguro. Por fim, a construção de *clusters* apresenta a última medida definida no projeto e, devido ao carácter inovador do mesmo, não se encontra apoiada em artigos científicos. No entanto, foram propostos limites mínimos e máximos para o *Score Final* obtido pelo cliente, por forma a permitir a atribuição, ou não, de um desconto associado a um prémio de um seguro.

5.2. Forças e limitações do estudo

Os principais pontos fortes deste projeto prendem-se com o carácter inovador do estudo aliado à pertinência em que o mesmo se desenvolve. Ora, até à data

não se encontra desenvolvido qualquer modelo de negócio que procure aplicar o estilo de consumo alimentar a uma bonificação associada ao prémio de um seguro de saúde ou vida. Este estudo procura instigar os primeiros trabalhos inerentes a este desenvolvimento. Paralelamente, dado o contexto atual do mercado, em que surgem novas empresas digitais nos serviços financeiros e que disponibilizam aos consumidores soluções mais simples, customizadas ou baratas, a criação de um seguro *data-driven* poderá permitir a criação de vantagens competitivas no mercado de seguros português, bem como a criação de um novo instrumento de fidelização associado ao seguro. De realçar o benefício associado ao ecossistema existente entre o cartão de pagamento da *FinTech* e o cartão de fidelização do Retalhista, que possibilita uma relação com o cliente, única e valiosa. Por último, aponta-se como uma força a representatividade das categorias de produto analisadas, cerca de 1 500 Unidades Base, bem como da amostra analisada, a aproximadamente 74 000 clientes.

Contudo, algumas limitações devem ser reparadas. Em primeiro lugar, dado o carácter inovador do estudo, e a disrupção do modelo de negócio que se pretende desenvolver, não foram encontradas referências que solidifiquem e justifiquem todas as metodologias escolhidas. Assim, a metodologia de segmentação baseada no padrão de consumo e que permite a aplicação de descontos associados ao prémio do seguro não está apoiada em artigos científicos, mas na distribuição dos resultados da amostra. Em segundo lugar, associada à construção do Modelo-SN reconhece-se uma lacuna que se prende com o facto de a informação nutricional não estar disponível para todos os produtos, bem como a segmentação em base de dados feita aos produtos pelo retalhista agregar, por vezes, artigos com algumas dissemelhanças nutricionais. Em terceiro lugar, a métrica *Score* Final incorpora no seu cálculo o total de transações realizadas pelo cliente, ponderado pelo *Score* Inicial e ajustado pelo

fator de conversão de cada produto e número de dependentes do cliente. No entanto, esta medida não considera o peso do produto para o desenvolvimento de uma dieta saudável. Concretizando melhor, considerando a distribuição alimentar proposta pela Roda dos Alimentos⁵, para a adoção de uma dieta equilibrada, o consumo de frutas e legumes deverá representar aproximadamente metade da alimentação de um indivíduo, comparativamente com o consumo de peixe, ovos e carne que é bastante menor. Complementarmente, a métrica não permite ajustar a amplitude de consumos no retalhista de cada cliente. De uma forma mais concreta, a métrica *Score Final* não distingue um cliente que adquire todo o seu cabaz alimentar no retalhista de um cliente que recorre ao retalhista para comprar determinados produtos, como frutas e legumes. Deste modo, um cliente que pretenda reduzir o seu prémio de seguro, poderá apenas fidelizar compras de alimentos mais saudáveis, enganando o modelo. Por último, ao apresentar um modelo de seguro que atribui descontos associados à prática de hábitos alimentares saudáveis podem ser levantadas algumas questões de ética visto que apenas os clientes com menor risco associado – ou seja, com padrões de consumo alimentares melhores – irão beneficiar de uma redução do seu prémio mensal.

5.3. Desenvolvimentos Futuros

Terminado o projeto, podemos identificar um conjunto de estudos e análises que deverão ser realizadas numa fase prévia ao lançamento deste novo modelo de negócio. Em primeiro lugar, importa analisar a recetividade dos consumidores em autorizar o acesso aos seus dados de consumo alimentares,

⁵ http://www.alimentacaosaudavel.dgs.pt/activeapp/wp-content/files_mf/1485170312CartazA313_med.pdf
(acedido a 27.02.2019)

bem como o seu interesse em aderir a um seguro com estas características – esta análise poderá ser alcançada com recurso a estudos do mercado ou questionários. Em segundo lugar, destaca-se o estudo da receptividade da seguradora parceira em reformular modelo de ponderação do risco, para que incorpore o Modelo-SN na sua análise – este ponto poderá ser resolvido com a realização de reuniões com seguradoras, onde a relação entre os hábitos alimentares e a saúde deverá ser analisada. Em terceiro lugar, será necessário desenvolver um modelo de *pricing* associado ao desconto a aplicar.

Por outro lado, identificam-se desenvolvimentos futuros que não afetam o lançamento do produto, mas podem condicionar a sua atratividade no mercado, fiabilidade e rentabilidade: i) reestruturação de determinadas categorias consumo, para que uma determinada categoria não inclua alimentos com uma grande disparidade nutricional; ii) incorporação de um conceito de *gamification* na aplicação móvel do cartão de pagamento, onde o cliente possa acompanhar e melhorar o seu perfil de consumo; iii) incorporação de inteligência artificial no Modelo-SN, no sentido de desenvolver um modelo robusto, evolutivo e dinâmico que consiga ir aprendendo à medida que surjam outras formas de classificar os alimentos, incluindo, por exemplo, técnicas de negociação avançadas entre diferentes modelos de classificação externos ao Modelo-SN; e iv) revisão da incorporação de dados biométricos (*wearables*) na ponderação do prémio associado ao seguro.

Referências Bibliográficas

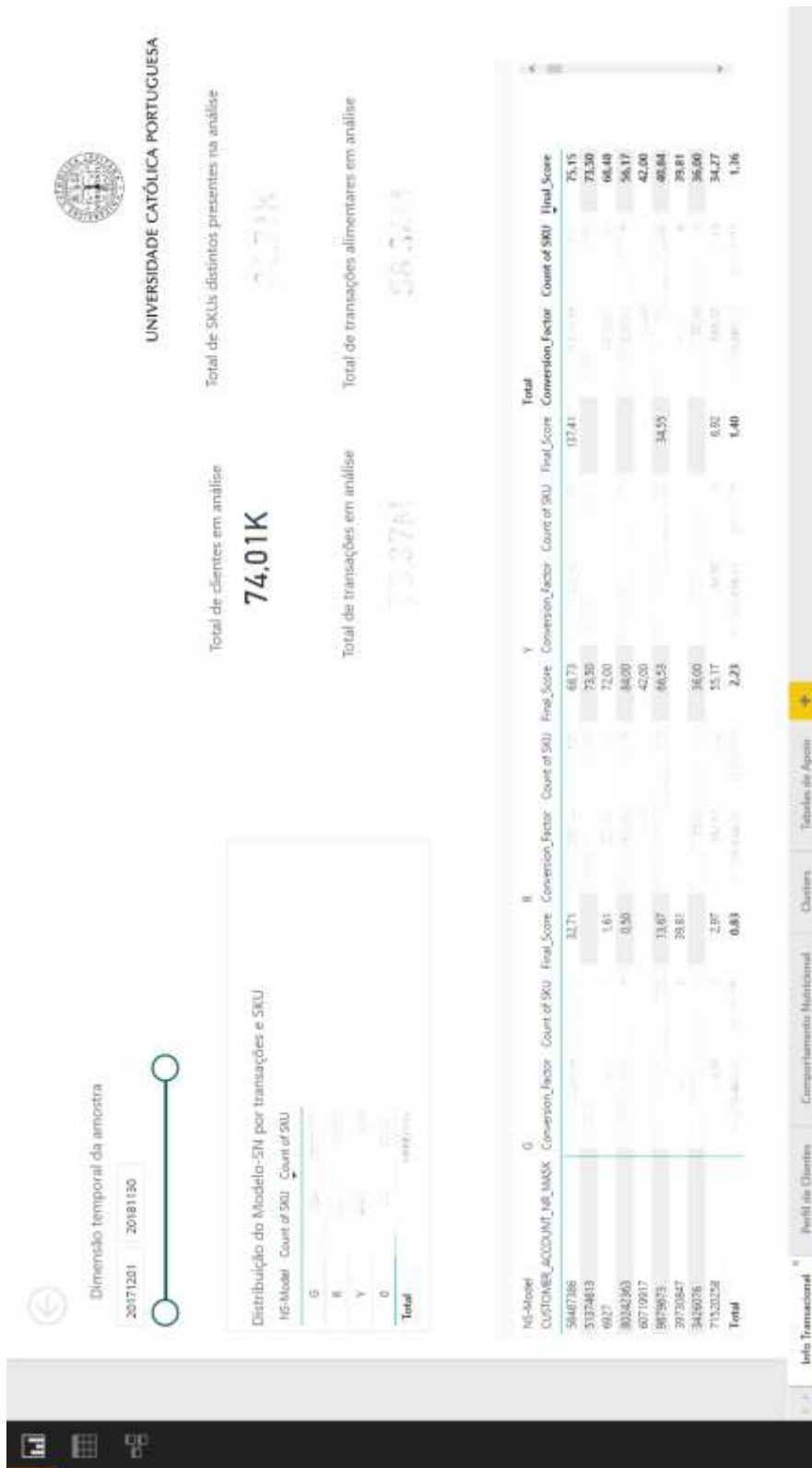
- Ahmad, A., Babar, M., Din, S., Khalid, S., Ullah, M. M., Paul, A., ... Min-Allah, N. (2019). Socio-cyber network: The potential of cyber-physical system to define human behaviors using big data analytics. *Future Generation Computer Systems*, 92, 868–878. <https://doi.org/10.1016/j.future.2017.12.027>
- Arechavala, T., Contente, X., Pérez-Giménez, A., Bartoll, X., Sànchez-Martínez, F., & López, M. J. (2016). Frecuencias de consumo de alimentos y exceso de peso en adolescentes en un contexto de crisis económica en Barcelona (España). *Gaceta Sanitaria*, 30(6), 457–463. <https://doi.org/10.1016/j.gaceta.2016.03.013>
- Boyd, D., & Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication & Society*, 15(5), 662–679. <https://doi.org/10.1080/1369118x.2012.678878>
- Coll, S. (2013). Consumption as biopower: Governing bodies with loyalty cards. *Journal of Consumer Culture*, 13(3), 201–220. <https://doi.org/10.1177/1469540513480159>
- Conner, T. S., Thompson, L. M., Knight, R. L., Flett, J. A. M., Richardson, A. C., & Brookie, K. L. (2017). The role of personality traits in young adult fruit and vegetable consumption. *Frontiers in Psychology*, 8(FEB), 1–11. <https://doi.org/10.3389/fpsyg.2017.00119>
- Costa-Tutusaus, L., & Guerra-Balic, M. (2016). Relationship between healthy lifestyle and sociodemographic factors in adolescents in Catalonia: Application of VISA-TEEN questionnaire. *PLoS ONE*, 11(9), 1–19. <https://doi.org/10.1371/journal.pone.0163381>
- Dawson, A. (2016). The economic essentials of digital strategy(changes the nature of supply, demand, or both). *McKinsey Quarterly*, (March).

- Dunford, E. K., Poti, J. M., Xavier, D., Webster, J. L., & Taillie, L. S. (2017). Color-coded front-of-pack nutrition labels—an option for US packaged foods? *Nutrients*, 9(5), 1–10. <https://doi.org/10.3390/nu9050480>
- Egnell, M., Talati, Z., Hercberg, S., Pettigrew, S., & Julia, C. (2018). Objective understanding of front-of-package nutrition labels: An international comparative experimental study across 12 countries. *Nutrients*, 10(10). <https://doi.org/10.3390/nu10101542>
- Food Standards Agency. (2007). Front of Pack Nutritional Signpost Labelling Technical Guidance, (1), 1–12. Retrieved from https://www.foodwatch.org/fileadmin/Themen/Ampelkennzeichnung/guidance_ampel_issue_1_januar_2007.pdf
- Gölzer, P., & Fritzsche, A. (2017). Data-driven operations management: organisational implications of the digital transformation in industrial practice. *Production Planning and Control*, 28(16), 1332–1343. <https://doi.org/10.1080/09537287.2017.1375148>
- Gundlach, K. H. H. (2017). The Impact of Startup Entry on the Innovativeness of Incumbents: Evidence from the Insurance Industry Dissertation submitted in partial fulfilment of requirements for the MSc in Management – Specialization Strategy and Entrepreneurship, at the Universidad, (June). Retrieved from [https://repositorio.ucp.pt/bitstream/10400.14/22699/1/Kilian Gundlach - The Impact of Startup Entry on the Innovativeness of Incumbents - Evidence from the Insurance Industry.pdf](https://repositorio.ucp.pt/bitstream/10400.14/22699/1/Kilian%20Gundlach%20-%20The%20Impact%20of%20Startup%20Entry%20on%20the%20Innovativeness%20of%20Incumbents%20-%20Evidence%20from%20the%20Insurance%20Industry.pdf)
- Hess, T., Matt, C., Wiesböck, F., & Benlian, A. (2016). Options for Formulating a Digital Transformation Strategy. *MIS Quarterly Executive*, 2016(June), 2–3.
- Holubcikova, J., Kolarcik, P., Madarasova Geckova, A., Van Dijk, J. P., & Reijneveld, S. A. (2016). Lack of parental rule-setting on eating is associated with a wide range of adolescent unhealthy eating behaviour both for boys and girls. *BMC Public Health*, 16(1). <https://doi.org/10.1186/s12889-016-3002->

- Meyer-Waarden, L. (2008). The influence of loyalty programme membership on customer purchase behaviour. *European Journal of Marketing*, 42(1–2), 87–114. <https://doi.org/10.1108/03090560810840925>
- Richi, E. B., Baumer, B., Conrad, B., Darioli, R., Schmid, A., & Keller, U. (2015). Health risks associated with meat consumption: A review of epidemiological studies. *International Journal for Vitamin and Nutrition Research*, 85(1–2), 70–78. <https://doi.org/10.1024/0300-9831/a000224>
- Roedder, N., Dauer, D., Laubis, K., Karaenke, P., & Weinhardt, C. (2016). The digital transformation and smart data analytics: An overview of enabling developments and application areas C3 - Proceedings - 2016 IEEE International Conference on Big Data, Big Data 2016, 2795–2802. <https://doi.org/10.1109/BigData.2016.7840927>
- Russom, P., & Org, T. (2011). *BIG DATA ANALYTICS FOURTH QUARTER 2011 TDWI BEST PRACTICES REPORT Introduction to Big Data Analytics. Definition from WhatIs.com.* Retrieved from <https://searchbusinessanalytics.techtarget.com/definition/big-data-analytics>
- Santos, S. (2018). *TRADITIONAL AND CHALLENGER BANKS IN UK: COMPARISON IN TERMS OF CUSTOMER VALUE.* Retrieved from <http://hdl.handle.net/10400.14/25884>
- Saunders, Mark; Lewis, Philip; Thornhill, A. (2011). *Research methods for business students.*
- Stoekli, E., Dremel, C., & Uebernickel, F. (2018). Exploring characteristics and transformational capabilities of InsurTech innovations to understand insurance value creation in a digital world. *Electronic Markets*, 28(3), 287–305. <https://doi.org/10.1007/s12525-018-0304-7>
- Teece, D. J. (2010). Business models, business strategy and innovation. *Long Range Planning*, 43(2–3), 172–194. <https://doi.org/10.1016/j.lrp.2009.07.003>

- Thorndike, A. N., Franckle, R. L., Polacsek, M., Boulos, R., Greene, J. C., Moran, A., ... Blue, D. J. (2017). A Supermarket Double-Dollar Incentive Program Increases Purchases of Fresh Fruits and Vegetables Among Low-Income Families With Children: The Healthy Double Study. *Journal of Nutrition Education and Behavior*, 50(3), 217–228.e1. <https://doi.org/10.1016/j.jneb.2017.09.013>
- Trichopoulou, A., Bamia, C., & Trichopoulos, D. (2009). Anatomy of health effects of Mediterranean diet: Greek EPIC prospective cohort study. *BMJ (Online)*, 339(7711), 26–28. <https://doi.org/10.1136/bmj.b2337>
- Voráčová, J., Sigmund, E., Sigmundová, D., & Kalman, M. (2016). Family affluence and the eating habits of 11-to 15-year-old Czech adolescents: HBSC 2002 and 2014. *International Journal of Environmental Research and Public Health*, 13(10), 23–25. <https://doi.org/10.3390/ijerph13101034>

Apêndice II – Ilustração da amostra em *Power BI Desktop*





Count of CUSTOMER_ACCOUNT_NIL_MASK by GENDER

GENDEER F M



Count of CUSTOMER_ACCOUNT_NIL_MASK by NEW_ZAL

NEW_ZAL 1 2 3 4 5



UNIVERSIDADE CATÓLICA PORTUGUESA

Count of CUSTOMER_ACCOUNT_NIL_MASK by SEG_LIFE_STYLE_KEY

SEG_LIFE_STYLE_KEY



Count of CUSTOMER_ACCOUNT_NIL_MASK by Seg_Age

Seg_Age

0-18 19-25 26-35 36-45 46-55 56-65 65 ou + Sem Valor



Info Transaccional

Perfil de Clientes

Comportamento Relacional

Clusters

Teclas de Apoio

Código Semáforo Nutricional
NS-Model

- Select All
- Blank
- G
- G
- R
- V

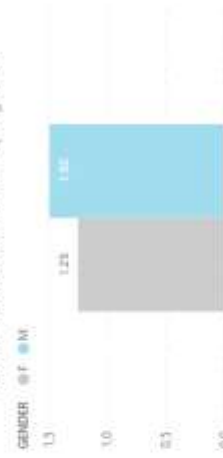
Média do Score Final do Modelo na Amostra



Distribuição do Modelo do Semáforo Nutricional na Amostra



Hábitos de consumo alimentares por género



Hábitos de consumo alimentares por Segmento



Tabela de Apoio

Clusters

Comportamento Nutricional

Perfil de Clientes

Análise Transacional



10 clusters

Select All	Cluster4
(Blank)	Cluster1
Cluster1	Cluster6
Cluster10	Cluster7
Cluster2	Cluster8
Cluster3	Cluster9

