

ISABELLA BARRAGAN RACHED

**REVISÃO DA SEGMENTAÇÃO DE LOJAS E DO PROCESSO DE PREVISÃO DE  
DEMANDA EM UMA EMPRESA VAREJISTA**

São Paulo

2019



ISABELLA BARRAGAN RACHED

**REVISÃO DA SEGMENTAÇÃO DE LOJAS E DO PROCESSO DE PREVISÃO DE  
DEMANDA EM UMA EMPRESA VAREJISTA**

Trabalho de Formatura apresentado à  
Escola Politécnica da Universidade de  
São Paulo para obtenção do diploma de  
Engenheira de Produção

São Paulo

2019



ISABELLA BARRAGAN RACHED

**REVISÃO DA SEGMENTAÇÃO DE LOJAS E DO PROCESSO DE PREVISÃO DE  
DEMANDA EM UMA EMPRESA VAREJISTA**

Trabalho de Formatura apresentado à  
Escola Politécnica da Universidade de  
São Paulo para obtenção do diploma de  
Engenheiro de Produção

Orientador: Prof. Dr. Mario Sergio  
Salerno

São Paulo  
2019

### Catálogo-na-publicação

Rached, Isabella Barragan

Revisão da segmentação de lojas e do processo de previsão de demanda em uma empresa varejista / I. B. Rached -- São Paulo, 2019.

111 p.

Trabalho de Formatura - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Produção.

1.Previsão de demanda 2.Segmentação de lojas 3.Varejo I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Produção II.t.

## AGRADECIMENTOS

Em primeiro lugar, gostaria de agradecer aos meus pais pelo apoio inabalável e amor incondicional ao longo da minha vida. Nenhuma conquista seria possível não fosse o incentivo de vocês, que acreditaram em mim mesmo quando eu não podia e me ajudaram a superar os momentos mais difíceis. Obrigada por nunca me deixarem desistir, por me presentear com a melhor família que eu poderia pedir e por serem exemplos de dedicação, integridade e perseverança – sem perder a ternura. Por essas e tantas outras razões, me sinto profundamente privilegiada de ser sua filha.

Ao meu irmão, leal companheiro, que me entende tanto apesar de sermos tão diferentes. Se hoje me formo engenheira, foi em parte pelo orgulho de você que me levou a seguir seus passos. Às minhas tias, primas, padrinhos e meus saudosos avós cuja confiança e incentivo permitiram que eu sonhasse alto e conquistasse grandes objetivos.

Ao Pedro, por estar ao meu lado nesse último ano e me mostrar que nenhum sonho é grande demais. Obrigada por confiar em mim, em todos os sentidos, e me fazer querer ser uma pessoa melhor.

À Gabriela, Maria Júlia, Felipe, Bruno e Thomas, que me acompanham há mais de quinze anos. Obrigada por fazerem parte dos melhores momentos da minha vida, por compartilharem as suas comigo e me motivarem desde sempre. Por me tocarem com sua determinação e doçura. Não existem palavras para agradecer a amizade e parceria dos cinco.

Às minhas melhores amigas, que me inspiram a ser melhor e mais dedicada todos os dias, na tentativa de me tornar um pouco mais parecida com vocês; que sorte a minha serem tantas que não consigo citar. Obrigada por me estimularem, me permitirem fazer parte da vida de cada uma e por estarmos juntas para tudo o que der e vier. Não estaria aqui hoje se não tivesse vocês comigo ao longo do caminho.

Aos meus queridos politécnicos, sem os quais tudo teria sido em vão. Ao Gabriel, por ser o melhor amigo em qualquer circunstância, pela sua genialidade que me orgulha e por estar ao meu lado quando mais precisei; que eu possa fazer o mesmo por você. Aos meus Andrés e Lucas, por estarem comigo desde o início dessa jornada, pelo apoio mútuo e a amizade que já se provou eterna. À Lina, Alice e Luiza, ao Joel, Victor Hugo, Mario, Lucas, Fernando, Rodrigo, Pedro e Gustavo por compartilharem os momentos mais felizes da minha graduação. Ao Mateus, meu amigo-irmão, pelo companheirismo, apoio e por ser quem você é.

Ao CAEP, por ter sido minha segunda casa, à Cris e ao Osni pelo infalível sorriso no rosto, o cuidado e carinho ao longo dos meus anos de graduação.

À Flávia, minha eterna *roomie*. Obrigada por abrir meus olhos para tantas reflexões, por me ajudar a me tornar a mulher que sou hoje e por compartilharmos os momentos mais especiais do nosso intercâmbio – e das nossas vidas desde então. Ao Lucca e Ariel, que se tornaram minha família longe de casa e nunca deixarão de ser meus irmãos. Obrigada por tanta parceria, intimidade, risadas e pelos momentos mais especiais. À Carol, minha amiga mais antiga, que sempre se fez presente; e à nossa grande família.

Aos tantos professores e professoras que me ajudaram a chegar até aqui e me tornar uma engenheira politécnica. Àqueles que me influenciaram e ensinaram desde cedo, despertando a vontade de aprender e crescer por meio do conhecimento. Em especial ao meu orientador, Prof. Dr. Mario Sergio Salerno, pelo apoio, conselhos e paciência durante o longo processo de realização desse projeto.

À Taise e ao Victor, por me ensinarem tanto e confiarem na minha capacidade. Ao restante do nosso time, pela parceria nas longas horas de trabalho, pela diversão que as fez passar mais rápido e por todos os aprendizados.

Finalmente, à Helena, a estrela mais brilhante, para quem dedico essa e todas as conquistas que vier a realizar; mesmo certa de que nenhuma será o bastante para retribuir o privilégio de ter sua amizade. Obrigada por ter me escolhido como amiga e confidente pelo eterno tempo que tivemos juntas. Te vejo em cada pôr do sol, sabendo que nenhum vai superar os que presenciei com você ao meu lado.



*“Take the sourest lemon life has to offer  
and turn it into something resembling  
lemonade”*



## RESUMO

O projeto foi desenvolvido em uma empresa varejista de bens de consumo, por meio de uma parceria dos times internos com um time da consultoria estratégica Bain & Company, do qual a autora fazia parte.

A falta de segmentação adequada das lojas da empresa gerava uma má gestão do sortimento delas, fazendo com que fosse oferecida uma variedade excessiva de produtos em toda a rede. Com isso, as lojas de menor área ficavam lotadas demais para expor todos os produtos e a experiência dos clientes era prejudicada - tanto pela lotação da área da exposição quanto pela falta de direcionamento do portfólio oferecido. Além disso, a baixa acurácia dos métodos de previsão de demanda utilizados levava a uma produção que não condizia com a demanda real, faltando produtos de alta demanda e sobrando outros ao longo da campanha.

O objetivo do projeto foi segmentar as lojas da Varejo SA de forma a redefinir o sortimento de forma personalizada e rever o processo de previsão de demanda utilizado para atualização do planejamento da produção. Como resultado, foi criado um procedimento de segmentação que pode ser revisto com a expansão da rede de lojas e uma metodologia simples e clara para definição do sortimento de cada segmento. Outros resultados foram a criação de um processo de previsão de demanda melhor definido, assim como um modelo quantitativo de previsão que apresenta uma redução de quase 80% nos erros em relação ao método anteriormente utilizado. Todos os produtos finais mencionados foram aprovados e implementados pela empresa ao longo ou ao final do projeto, sendo seus resultados reais ou esperados apresentados no decorrer desse trabalho.

**Palavras-chave:** Segmentação de lojas. Definição de sortimento. Previsão de demanda.



## ABSTRACT

The project was developed in a consumer goods retail company, through a partnership of internal teams with a team of the strategic consulting firm Bain & Company, in which the author participated.

The lack of proper segmentation of the company's stores generated poor management of their assortment, offering an excessive variety of products throughout the network. Therefore, the smaller stores were too crowded because all products were exposed, and customer experience was impaired – both by the crowded exposition space and the lack of focus of the available portfolio. In addition, the low accuracy of the demand forecasting methods used generated errors in production planning, which could be deeply harmful to the company. The deregulated forecast led to a production that did not match real demand, in a way that stores ran out of high-demand products and overstocked others.

The goal of the project was to segment Varejo SA stores in order to redefine the assortment in a personalized way and review the demand forecasting process used to update production planning. As a result, a segmentation procedure was created, which can be reviewed with the expansion of the store network, as well as a simple and clear methodology for defining the assortment of each segment. Other results were the creation of a well-defined demand forecasting process and a quantitative model to be used by the planning team that reduced errors by almost 80% when compared to the previous one. All the project's products were approved and implemented by the company throughout the project or at the end of it. Therefore, their results, either real or expected, are presented in this paper.

**Keywords:** Store segmentation. Assortment definition. Demand forecasting.



## LISTA DE FIGURAS

Figura 1: Escritórios da Bain & Company .....	18
Figura 2: Pirâmide de produtos da Varejo SA.....	21
Figura 3: métodos quantitativos usados no marketing .....	27
Figura 4: algoritmo de clusterização hierárquica aglomerativa.....	29
Figura 5: exemplo de dendograma .....	29
Figura 6: algoritmo de <i>K-means</i> .....	30
Figura 7: elementos envolvidos no cálculo de $s(i)$ , $i \in A$ .....	33
Figura 8: hierarquia de decisões de sortimento .....	35
Figura 9: planejamento e controle da produção.....	38
Figura 10: previsão da demanda no planejamento da produção.....	41
Figura 11: exemplos de métodos de previsão de demanda .....	50
Figura 12: benchmark de segmentação de canais.....	62
Figura 13: etapas e atividades de planejamento de campanha .....	80
Figura 14: etapas do processo de previsão proposto .....	81
Figura 15: divisão dos dados históricos para construção do modelo .....	94





## LISTA DE GRÁFICOS

Gráfico 1: gráfico de projeção do número de lojas da Varejo SA, base 100.....	19
Gráfico 2: número de itens, volume vendido e receita bruta por tipo de SKU .....	22
Gráfico 3: projeção da receita bruta da Varejo SA no Brasil, em base 100 .....	23
Gráfico 4: demanda no tempo .....	44
Gráfico 5: gráfico de demanda por tempo para duas séries.....	46
Gráfico 6: gráfico de sazonalidade teórico .....	47
Gráfico 7: correlograma de uma série histórica com sazonalidade anual .....	48
Gráfico 8: gráfico acumulado de receita bruta por SKU, Jun/18-Mai/19 .....	63
Gráfico 9: gráfico de densidade por área de lojas selecionadas da Varejo SA .....	64
Gráfico 10: gráfico de receita bruta (%) por categoria das lojas da Varejo SA, 2018 .....	65
Gráfico 11: gráfico de receita bruta (%) por linha da principal categoria da Varejo SA .....	66
Gráfico 12: distribuição da silhueta para 3 clusters.....	68
Gráfico 13: representação gráfica dos cluster obtidos.....	69
Gráfico 14: representação gráfica dos clusters finais .....	71
Gráfico 15: relevância dos grupos de produtos definidos em número de itens e receita bruta do varejo .....	74
Gráfico 16: gráfico de ruptura e cobertura de estoque, jun/18 a mai/19 .....	77
Gráfico 17: gráfico de demanda realizada e prevista, ago/18 a jul/19 .....	78
Gráfico 18: relevância das famílias no total de itens vendidos e receita bruta, 2017 e 2018...83	
Gráfico 19: gráfico de sazonalidade da demanda de bolsas neutras, Jan/2017-Ago/2019.....	86
Gráfico 20: demanda por campanha, Jan/2017-Ago/2019 .....	87
Gráfico 21: gráfico de sazonalidade da demanda de bolsas coloridas, Jan/2017-Ago/2019....	87
Gráfico 22: demanda por campanha, Jan/2017-Ago/2019 .....	88
Gráfico 23: gráfico de sazonalidade da demanda de cintos femininos, Jan/2017-Ago/2019...89	
Gráfico 24: demanda por campanha, Jan/2017-Ago/2019 .....	89
Gráfico 25: gráfico de sazonalidade da demanda de cintos femininos, Jan/2017-Ago/2019...90	
Gráfico 26: demanda por campanha, Jan/2017-Ago/2019 .....	90
Gráfico 27: gráfico de sazonalidade da demanda de carteiras femininas, Jan/2017-Ago/201991	
Gráfico 28: demanda por campanha, Jan/2017-Ago/2019 .....	92
Gráfico 29: Previsão da demanda de bolsas de cores neutras .....	96
Gráfico 30: Previsão da demanda de bolsas coloridas .....	97
Gráfico 31: Previsão da demanda de cintos femininos.....	98

Gráfico 32: Previsão da demanda de cintos masculinos .....	99
Gráfico 33: Previsão da demanda de carteiras femininas .....	100
Gráfico 34: correlograma da demanda de bolsas neutras (A) e coloridas (B) .....	110
Gráfico 35: correlograma da demanda de cintos femininos (A) e masculinos (B).....	111
Gráfico 36: correlograma da demanda de carteiras femininas.....	111

## LISTA DE TABELAS

Tabela 1: interpretação sugerida dos coeficientes de silhueta.....	34
Tabela 2: coeficiente de silhueta das clusterizações testadas .....	68
Tabela 3: centroides dos clusters .....	69
Tabela 4: erros percentuais da previsão do software, campanha 2018.09 a 2019.08.....	78
Tabela 5: erros das previsões para bolsas de cores neutras .....	97
Tabela 6: erro das previsões para bolsas coloridas.....	98
Tabela 7: erros das previsões para cintos femininos .....	99
Tabela 8: erros das previsões para cintos masculinos .....	100
Tabela 9: erros das previsões para carteiras femininas.....	101



## SUMÁRIO

1	INTRODUÇÃO .....	16
1.1	Descrição do estágio .....	17
1.1.1	A Bain & Company .....	17
1.2	A empresa estudada .....	18
1.2.1	Breve apresentação .....	18
1.2.2	Rede de lojas.....	19
1.2.3	Estrutura de produtos .....	20
1.3	O Planejamento Comercial da Varejo SA .....	22
1.4	Relevância do trabalho.....	23
2	FUNDAMENTAÇÃO TEÓRICA SOBRE SEGMENTAÇÃO DE CANAIS E SORTIMENTO DE PRODUTOS .....	24
2.1	Motivações para segmentar mercados .....	24
2.2	Variáveis para segmentação.....	25
2.3	Métodos de clusterização .....	26
2.3.1	Métodos hierárquicos aglomerativos.....	28
2.3.2	Método de <i>K-means</i> .....	30
2.4	Validação da clusterização.....	32
2.5	Sortimento de produto .....	34
3	FUNDAMENTAÇÃO TEÓRICA SOBRE PREVISÃO DE DEMANDA.....	38
3.1	Sistema de Planejamento e Controle da Produção (PCP).....	38
3.2	Processo de S&OP .....	39
3.3	Importância de prever .....	39
3.4	Previsão e Planejamento .....	40
3.5	Princípios de previsão .....	41
3.6	Procedimento para realização da previsão .....	42
3.7	Características da demanda.....	43

3.7.1	Inspeção gráfica .....	45
3.8	Técnicas de previsão .....	49
3.8.1	Métodos quantitativos de projeção .....	50
3.8.2	Ajustes e transformações .....	55
3.9	Acurácia da previsão .....	56
3.10	Seleção dos métodos de previsão .....	59
4	SEGMENTAÇÃO DE LOJAS E DEFINIÇÃO DO SORTIMENTO DA VAREJO SA	60
4.1	Propósito da segmentação .....	62
4.2	Exploração das variáveis a serem consideradas .....	64
4.3	Escolha do método de clusterização.....	66
4.4	Resultados e validação da clusterização.....	67
4.5	Características dos segmentos obtidos .....	69
4.6	Redefinição do sortimento.....	71
4.7	Resultados .....	73
5	PREVISÃO DE DEMANDA NA VAREJO SA.....	76
5.1	Propósito da previsão .....	79
5.2	Procedimento de previsão proposto .....	80
5.3	Nível de agregação da previsão.....	81
5.4	Coleta de dados .....	83
5.5	Análise preliminar .....	85
5.6	Escolha dos métodos utilizados.....	92
5.7	Estrutura do modelo .....	93
5.8	Avaliação dos resultados .....	95
5.9	Implementação do modelo .....	101
6	CONCLUSÃO.....	76
6.1	Síntese e resultados .....	104
6.2	Análise crítica.....	105

6.3	Desdobramentos.....	106
7	REFERÊNCIAS BIBLIOGRÁFICAS .....	108
	ANEXO A: AUTOCORRELAÇÃO DAS SÉRIES HISTÓRICAS .....	110





## 1 INTRODUÇÃO

É inegável a importância de processos de planejamento robustos no varejo, desde a previsão da demanda total por período até a alocação de produtos loja a loja. Do ponto de vista do consumidor, é altamente desejável que a loja de sua preferência tenha disponíveis os produtos que procura quando a visita, evitando desgastes e viagens desnecessárias. Para a empresa, é essencial ter estimativas de vendas confiáveis que permitam tomar as decisões necessárias do ponto de vista estratégico, tático e operacional de maneira bem informada. O planejamento de recursos, programação da produção, gestão de estoques e logística além de calendário comercial são alguns exemplos de atividades afetadas pelo planejamento comercial de uma empresa.

Com lojas abastecidas pelos produtos certos – que o consumidor procura – na quantidade certa, é possível evitar o custo extra de gestão de estoques superlotados, insatisfação dos clientes pela falta de itens, devoluções por excesso de produtos ou envios emergenciais de mercadorias faltantes. A importância da gestão de estoques é evidente, mas a eficiência desse processo depende de um passo anterior no planejamento: a previsão da demanda.

É necessário prever a demanda dos próximos períodos para que se possa planejar a produção, reposição e envio de produtos, a alavancagem de itens com baixo giro por meio de estratégias comerciais e outras iniciativas importantes.

O objetivo do presente estudo, portanto, é de reformular o processo de planejamento de demanda da Varejo SA, fornecendo uma metodologia que entregue uma previsão acurada por meio de ferramentas de simples utilização e com output claro. Como a empresa já operava em outros canais de venda e tem um amplo portfólio de itens, foi necessário primeiramente definir o sortimento que deveria ser enviado para as lojas, evitando acúmulo de produtos que não tem vazão no varejo.

Para tanto, foi realizado um diagnóstico da área de planejamento comercial de forma a entender quais os maiores pontos de dor e possíveis alavancas de melhoria. Em seguida, trabalhou-se a segmentação das lojas de forma que os grupos resultantes pudessem ser usados para definição do sortimento adequado de produtos, que pode variar a depender de características das lojas posteriormente descritas. Com o resultado desse trabalho, foi finalmente definido o sortimento para cada segmento (ou *cluster*), de forma que fosse possível planejar a demanda de cada um deles. Por fim, se desenvolveu uma metodologia de previsão para programação de campanhas, que equilibra a robustez e acurácia com a simplicidade e

rapidez de aplicação, solucionando os principais pontos de dor de planejamento da empresa estudada.

## **1.1 Descrição do estágio**

A autora do presente trabalho realizou estágio na empresa de consultoria estratégica Bain & Company por um período de, ao todo, quase 2 anos. Durante 6 meses desenvolveu o estágio em uma área interna de apoio aos times de consultores, relacionada a obtenção de dados de mercado, pesquisas e materiais das empresas estudadas e seus concorrentes, além de informações socioeconômicas em geral.

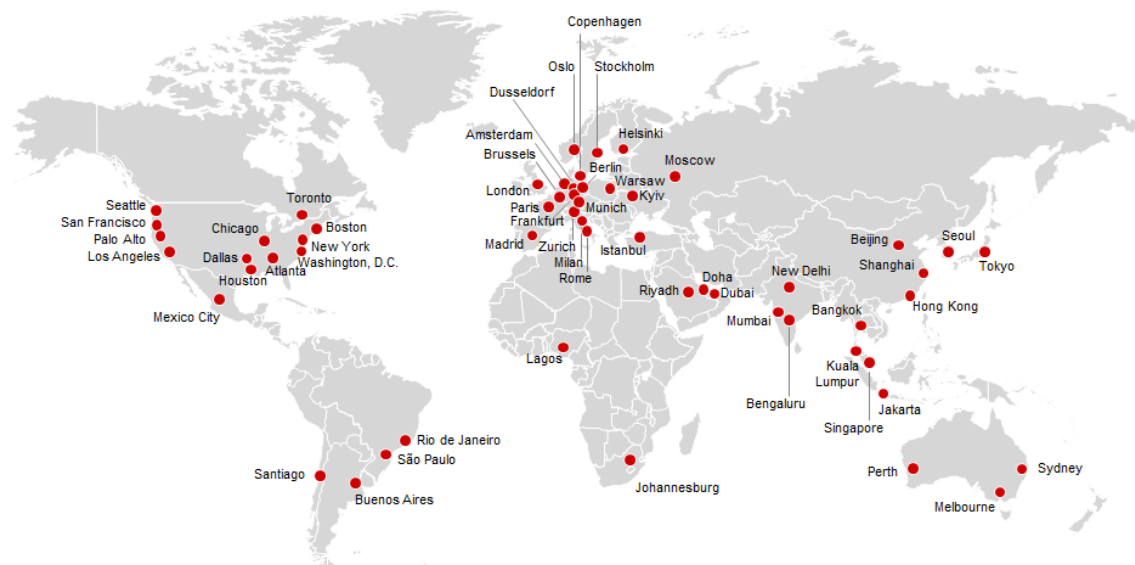
Após interrupção do contrato por um ano para realização de um aproveitamento de estudos na Universidade Politécnica de Madrid, a autora foi convidada e retomar o estágio na área de pesquisa da empresa e, no início de 2019, juntou-se ao time de consultoria. Ao longo do ano participou de dois projetos no mesmo grupo varejista de bens de consumo, abordado nesse trabalho.

### **1.1.1 A Bain & Company**

A Bain & Company é uma empresa global de consultoria estratégica, fundada e sediada nos Estados Unidos, na cidade de Boston. Presente em 37 países com 58 escritórios em diversas cidades, a empresa é uma das mais renomadas no ramo e atua em projetos nas mais diversas indústrias, serviços e capacidades.

A autora do trabalho em questão desenvolveu o estágio no escritório de São Paulo, enquanto realizava seus estudos na Escola Politécnica. Fundado em 1997, o escritório da capital possui atualmente mais de 500 funcionários e um amplo portfólio de clientes, sendo o principal da consultoria na região da América do Sul.

Figura 1: Escritórios da Bain &amp; Company



Fonte: material veiculado internamente pela Bain

Como anteriormente mencionado, a Bain atende uma ampla variedade de clientes e áreas de expertise, que incluem desde indústrias pesadas (agronegócio, infraestrutura, aviação, indústria química) até varejo, mídia, empresas de tecnologia e telecomunicação. O presente estudo foi realizado em um dos clientes da empresa, descrito ao longo do restante desse capítulo.

## 1.2 A empresa estudada

Nesse item será apresentada a empresa sobre a qual o trabalho em questão foi desenvolvido. Além de uma breve apresentação do histórico da companhia - que por questões de confidencialidade não poderá ser extremamente detalhada - será explicitada uma visão da estrutura de produtos comercializados pela empresa, necessária para a compreensão da delimitação do problema e desenvolvimento do projeto como um todo.

### 1.2.1 Breve apresentação

A Varejo SA é uma das maiores empresas de bens de consumo no país. Está em operação há mais de 40 anos por meio de diversos canais, desde a revenda em lojas multimarcas e especializadas até a venda direta por meio de consultoras associadas. Além de comercializar em diferentes canais, a empresa também é responsável pela produção dos itens vendidos,

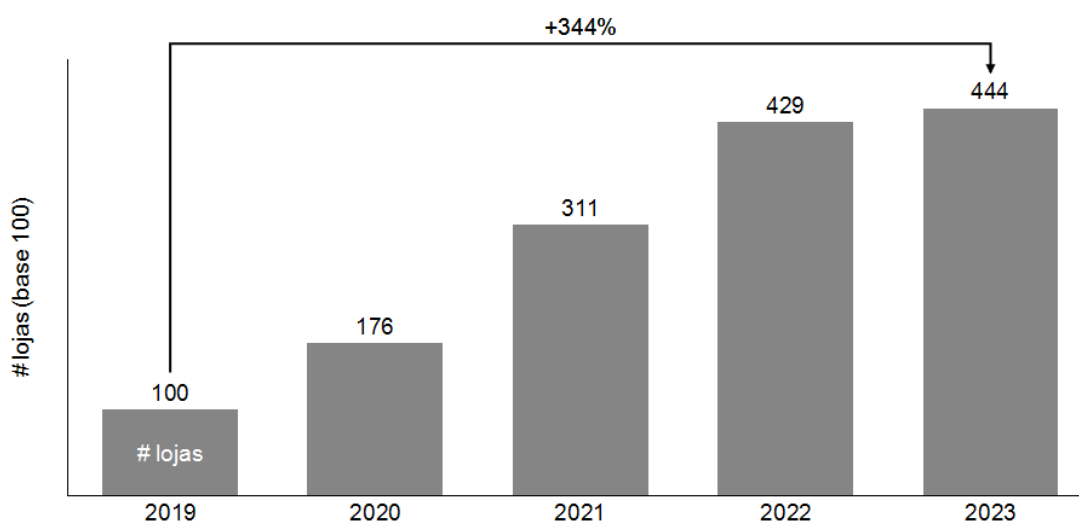
contando com uma fábrica no estado de São Paulo e centros de distribuição localizados pelo Brasil, sobretudo na região Sudeste.

Apesar de sua origem nacional, a empresa opera em outros países da América Latina e do mundo, tanto por meio da venda direta quanto, mais recentemente, pelas lojas próprias. A empresa entrou para o varejo há poucos anos, como parte da estratégia de expandir seu *footprint* e alcançar novos consumidores, inacessíveis por meio dos demais canais pelos quais opera. O canal representa, atualmente, pouco mais de 1% da receita da empresa no Brasil, o que evidencia seu estágio inicial.

### 1.2.2 Rede de lojas

A Varejo SA tem uma rede de lojas ainda reduzida, mas o plano estratégico da empresa prevê um aumento expressivo desse número nos próximos 5 anos, conforme apresenta o gráfico de projeção do número de lojas da empresa, construído com base nesse plano. As lojas hoje abertas têm um perfil razoavelmente similar em termos de localização, público e fluxo - sendo a maior parte localizada em shoppings de classe média ou alta, de alto fluxo e localizadas, sobretudo, na região sudeste.

Gráfico 1: gráfico de projeção do número de lojas da Varejo SA, base 100



Fonte: elaboração da autora a partir de material interno da Varejo SA

Futuramente, no entanto, o plano da empresa é que o perfil das lojas seja cada vez mais diversificado, sendo necessário tratar o sortimento da rede de forma personalizada para garantir

o sucesso dela. Para tanto, será descrito no capítulo 4 o estudo feito com base nas lojas atuais e ambições estratégicas da empresa de forma a definir segmentos de lojas de acordo com as dimensões relevantes, podendo por fim definir-se os produtos adequados a serem enviados para cada um deles.

### 1.2.3 Estrutura de produtos

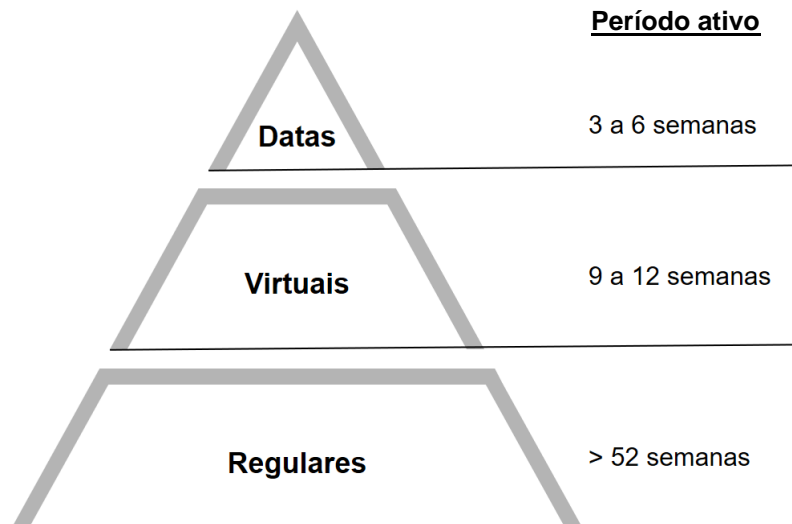
A empresa estudada possui uma ampla gama de produtos, podendo oferecer cerca de 1000 *Stock Keeping Units* (SKUs) simultaneamente em cada uma de suas lojas em determinadas épocas do ano.

O portfólio de produtos oferecidos na indústria em que a Varejo SA trabalha varia com o tempo não só por conta da descontinuação ou lançamento de itens, mas também pela sazonalidade que afeta o padrão de consumo ao longo do ano. A Varejo SA, portanto, trabalha com produtos que apresentam diversos períodos ativos – tempo durante o qual estão ativos e disponíveis para compra – diferentes, podendo ser divididos entre regulares, virtuais e kits de datas.

- Regulares: são produtos oferecidos ao longo do ano todo, independente da época ou de datas comemorativas. Seu período ativo costuma ser superior a 52 semanas, salvo exceções que são descontinuadas devido à baixa performance ou outros fatores externos. Dentre esses SKUs, existem os kits perenes, que são conjuntos de produtos regulares vendidos a um valor promocional, identificados pelo consumidor como mais presenteáveis do que um item isolado.
- Virtuais: são itens que entram e saem do portfólio disponível nas lojas sazonalmente. Podem ser, por exemplo, coleções de produtos desenvolvidos especialmente para uma estação. Esses itens costumam ficar ativos por períodos de 9 a 12 semanas e podem ser recorrentes, ativados todo o ano na mesma época, ou lançamentos únicos.
- Kits de datas: com período ativo mais curto que os virtuais, os kits de datas são lançados como estratégia de venda para uma campanha específica do calendário comercial, como Dia das Mães ou Natal. São combos de SKUs (regulares ou virtuais) montados especialmente para a venda promocional nos períodos em questão. Dessa forma,

costumam ficar disponíveis por apenas 3 a 6 semanas, a depender do feriado ao qual se destinam, sendo em seguida retirados do portfólio.

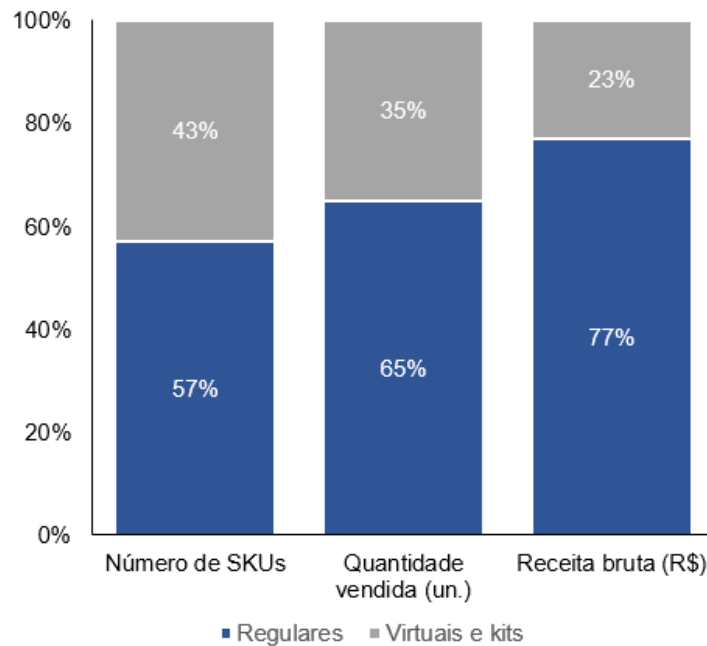
Figura 2: Pirâmide de produtos da Varejo SA



Fonte: elaboração da autora.

Quando se trata da relevância de cada um desses tipos de produtos, é esperado que os itens regulares sejam os mais demandados, devido à sua condição de maior perenidade em relação aos demais. No gráfico 2, se nota que os itens contínuos no portfólio representam mais da metade do total de SKUs, 65% da quantidade vendida ao longo do ano e quase 80% da receita bruta do varejo. Vale notar que a Varejo SA fabrica e comercializa produtos que tem uma margem de lucro muito semelhante e elevada, sendo indiferente compará-los pelo impacto na receita ou no lucro da operação.

Gráfico 2: número de itens, volume vendido e receita bruta por tipo de SKU



Fonte: elaborado pela autora a partir de dados da Varejo SA

### 1.3 O Planejamento Comercial da Varejo SA

A Varejo SA entrou no ramo do varejo há poucos anos, tendo um histórico reduzido de dados de venda e pouca experiência no canal, de forma que os processos específicos relacionados a ele ainda não foram completamente estruturados e testados. Dentre esses processos, diversos estão relacionados ao planejamento comercial, que é uma competência fundamental para o sucesso da empresa e atendimento da demanda dos clientes.

Mesmo na escala atual da operação, que consiste em poucas lojas, existem lacunas importantes que impactam a performance do negócio. Mais detalhes sobre o cenário atual e revisões necessárias serão apresentados nos capítulos 4 e 5, mas, de forma resumida, esses foram os principais problemas identificados:

- Quando da abertura das lojas, não foi realizado um estudo profundo de forma a definir o sortimento enviado para cada uma delas. Assim, existe uma oferta muito elevada de produtos que não vendem, ocupando espaço de exposição, estoque e reduzindo a produtividade das lojas;
- As metodologias de previsão de demanda utilizadas não são robustas, havendo frequentes erros no planejamento da produção e consequente excesso de estoque de

produtos de giro baixo e ruptura de estoque de itens *best sellers*, principalmente em feriados comerciais e datas comemorativas (Dia das Mães, Dia dos Pais, Natal etc.)

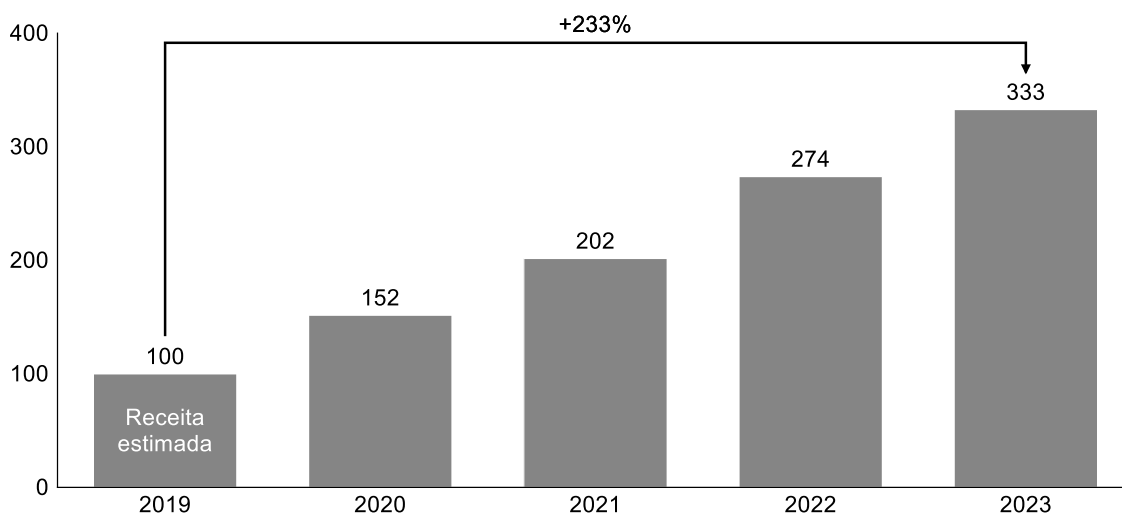
#### 1.4 Relevância do trabalho

A empresa em estudo tem planos bastante ambiciosos de crescimento. Segundo o plano estratégico do varejo para os próximos 5 anos, desenvolvido pela própria Varejo SA, é previsto um aumento de mais de quatro no número de lojas no Brasil, ao mesmo tempo em que se eleva a receita no país a uma taxa anual composta de mais de 40%, como mostra o gráfico 3.

Ainda segundo o plano estratégico elaborado pela empresa, uma das alavancas de crescimento é a adequação do portfólio oferecido nas lojas por meio da definição de agrupamentos de lojas que apresentem um comportamento de demanda e características de funcionamento semelhantes.

Dessa forma, é de extrema importância que a empresa ataque as questões problemáticas do planejamento comercial com a maior rapidez possível, dado que a falta de processos robustos será uma barreira para que o negócio escale de forma rentável, podendo colocar em risco o alcance dos objetivos desse plano estratégico.

Gráfico 3: projeção da receita bruta da Varejo SA no Brasil, em base 100



Fonte: elaboração da autora a partir de dados internos da Varejo SA



## **2 FUNDAMENTAÇÃO TEÓRICA SOBRE SEGMENTAÇÃO DE CANAIS E SORTIMENTO DE PRODUTOS**

### **2.1 Motivações para segmentar mercados**

A segmentação é o ato de dividir uma dada população (ou amostra dela) em grupos com significado (KOTLER; ARMSTRONG, 2016), cujos membros tenham características comuns entre si e diversas daqueles dos demais grupos. Essa divisão deveria fornecer uma base para tomada de decisões bem informadas no contexto organizacional, por meio da redução da amostra a segmentos coesos que podem ser tratados de uma mesma forma. De acordo com Blythe (2005), cada consumidor tem suas necessidades e desejos e, dada a impossibilidade de se atender cada um deles individualmente, faz sentido separá-los em grupos com características semelhantes.

No caso da segmentação de canais, não é diferente. Sendo os canais as lojas, no caso do varejo, sua segmentação consiste em dividi-las em grupos coesos e acionáveis, permitindo que a organização tome decisões estratégicas e operacionais com base neles. A divisão desses grupos pode ser feita por diversos critérios, sendo comumente baseada em aspectos socioeconômicos, geográficos e características físicas do estabelecimento.

São diversos os exemplos na literatura da aplicação do método de segmentação para planejamento da expansão da rede de lojas. Aplica-se, normalmente, na escolha da localização de novas unidades. No entanto, existem diversas outras aplicações relacionadas ao merchandising de uma empresa que podem ser auxiliadas pela divisão da rede em segmentos, como a definição do sortimento que deve variar conforme o perfil do consumidor e características espaciais das lojas da rede. Para Bermingham, Hernandez e Clark (2013), o valor agregado da segmentação de loja para as cadeias de varejo e serviço inclui a capacidade de adaptar marketing, merchandising e abordagens de publicidade por tipo de loja, agilizar a logística, delimitar territórios de vendas e avaliar o desempenho individual de lojas em comparação às suas similares. Blythe (2005) também apresenta algumas vantagens da segmentação de mercados, como a compreensão do perfil dos consumidores, concorrentes, alocação eficiente de recursos, planejamento de marketing estratégico e possibilidade de expansão do mercado.

## 2.2 Variáveis para segmentação

Blythe (2005) afirma que cada segmento definido deve cumprir certos requisitos para que a divisão seja eficaz, eficiente e alcance seus objetivos:

- Deve ser mensurável, de forma que se possa identificar os membros do segmento e sua quantidade;
- Deve ser acessível, ou seja, deve ser possível comunicar-se com o grupo e entregar o produto a ele;
- Deve ser substancial, grande o bastante para que valha à pena definir estratégias específicas para o segmento;
- Deve ser coerente, de forma que os usuários tenham de fato necessidades e padrões de consumo semelhantes;
- Deve ser estável, com membros e natureza relativamente constantes.

O processo de segmentação inicia-se com a seleção de variáveis significativas, a partir das quais o mercado possa ser diferenciado a ponto de definir grupos distintos. São muitos os estudos de caso que utilizam critérios diferentes para a segmentação dos clientes ou canais de venda de um varejista, sendo geralmente relacionados a características demográficas, hábitos de consumo dos clientes, características das lojas (como metragem, espaço de estoque, localização, fluxo, proximidade com centros comerciais etc.) e histórico de vendas. Segundo Blythe (2005), as principais bases para segmentação são:

- Geográfica: local, clima, topologia, etc. Por exemplo, os carros vendidos em locais de clima quente sempre têm ar condicionado potente, enquanto os comercializados em lugares de clima frio devem ter aquecedor;
- Psicográfica: baseada na personalidade dos indivíduos. Os seguros imobiliários, por exemplo, podem focar mais no segmento de pessoas que tem medo de crimes, desastres naturais etc.;
- Comportamental: essa abordagem foca nos benefícios, situação de uso, aplicações do produto e lealdade dos clientes. O mercado de carros, por exemplo, pode direcionar produtos aos usuários que utilizam o produto como simples meio de transporte urbano,

aqueles que dirigem carros esportivos como hobby ou que usam o automóvel para realizar trilhas, acampamentos e atividades ao ar livre;

- Demográfica: preocupa-se com a estrutura populacional em termos de idade, estilo de vida, fatores econômicos. O mercado de imóveis, por exemplo, pode ser segmentado em casas para aposentados, famílias com crianças, jovens cujo imóvel seria a primeira compra. É uma das abordagens mais utilizadas pela facilidade de obtenção de dados junto a órgãos públicos e eficiência que representam para alcançar as metas de segmentação dos mercados.

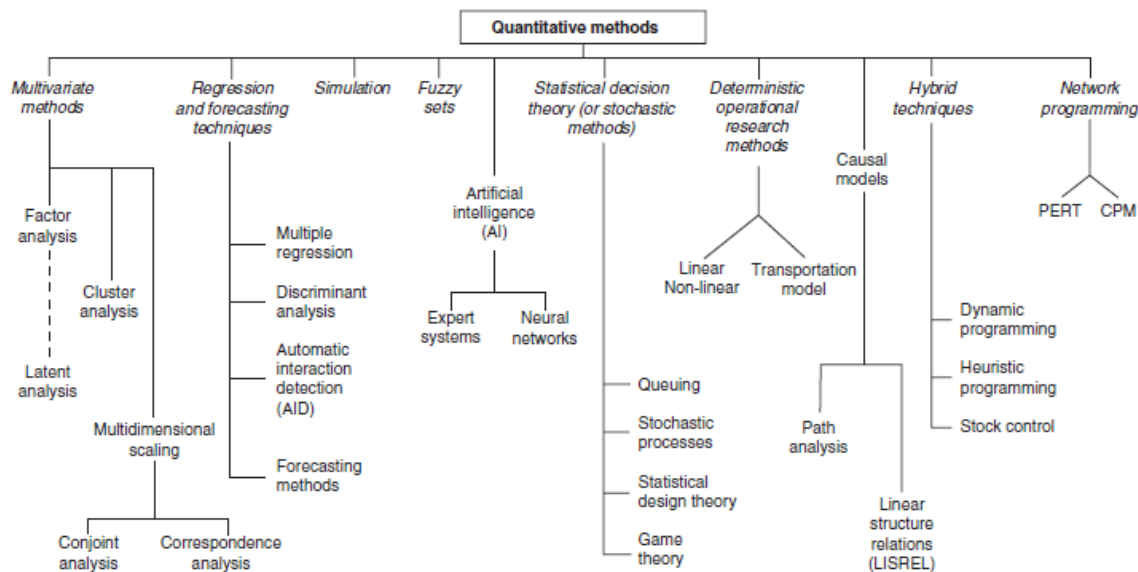
Uma segmentação confiável e acionável deve ser realizada com base em dados relevantes e seguir algum método quantitativo adequado. Um processo amplamente utilizado é o de clusterização, que costuma apresentar resultados robustos: grupos relevantes, suficientemente diferentes e que podem ser usados para a tomada de decisões ou análises a que se propõem. Segundo framework de Xu e Wunsch (2005), a clusterização deve seguir quatro passos:

- Seleção de fatores/variáveis a serem considerados
- Seleção e aplicação do algoritmo de clusterização
- Validação dos clusters obtidos
- Interpretação dos resultados

### **2.3 Métodos de clusterização**

São diversos os métodos quantitativos utilizados para análise de dados de marketing, os quais frequentemente tratam variáveis interdependentes como preços, oferta, distribuição, vendas. Moutinho (2003) apresenta diversos métodos, conforme o diagrama disposto na figura 3. Segundo Moutinho (2003) e Saunders (1980), o método de clusterização é o mais utilizado para segmentação de mercados e tomada de decisões estratégicas com base nos segmentos. A análise por clusters em sua variedade é a técnica mais utilizada para aplicações em serviços e varejo (BERMINGHAM; HERNANDEZ; CLARKE, 2013).

Figura 3: métodos quantitativos usados no marketing



Fonte: extraído de Moutinho (2003)

Existem diversos algoritmos de clusterização na literatura, sendo criados novos modelos frequentemente, já que não existe solução ideal para cada caso; diferentes modelos podem ser implementados e testados, a depender dos dados disponíveis, qualidade das ferramentas utilizadas, objetivo da clusterização, entre outros fatores.

Apesar da diversidade de métodos disponível, é consenso a divisão entre dois principais tipos: métodos hierárquicos e não hierárquicos. Segundo Dolničar (2003) os dois tipos são igualmente populares nos estudos relevantes: 46% deles utilizam métodos particionais e 44% hierárquicos.

- Métodos hierárquicos: técnicas nas quais os dados são particionados sucessivamente, produzindo uma representação hierárquica dos agrupamentos (BERMINGHAM; HERNANDEZ; CLARKE, 2013). O objetivo é encontrar a divisão que maximize as similaridades entre itens do mesmo agrupamento e aumente a diferença entre grupos distintos. Podem ser de dois tipos:
  - Aglomerativos: iniciados com cada ponto formando seu próprio cluster. Eles são gradualmente unidos até gerar grupos com um alto grau de similaridade;
  - Divisivos: iniciados com um único cluster que contém todos os pontos, o qual vai sendo dividido gradualmente em agrupamentos menores.

- Métodos não hierárquicos: também conhecidos como particionais, esses métodos procuram dividir a amostra de forma a minimizar ou maximizar um critério definido previamente, como a distância entre os pontos utilizados. O algoritmo de *K-means* é o mais amplamente mencionado na literatura relevante, sendo utilizado em 76% dos estudos que realizam uma clusterização por método particional (DOLNIČAR, 2003)

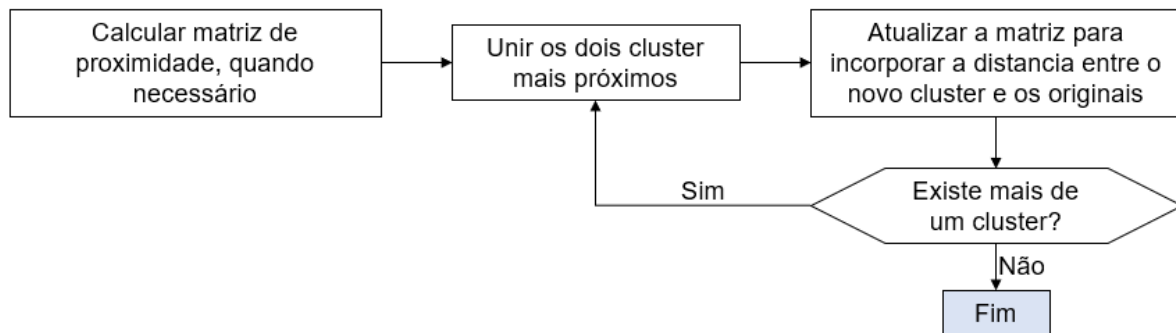
A literatura é clara e extensiva quanto às diferenças, vantagens e desvantagens de cada um dos métodos mencionados. Os hierárquicos são muito suscetíveis a outliers, já que uma vez designado a um cluster, um objeto não é revisto – assim, o modelo é incapaz de corrigir classificações malfeitas. Além disso, são métodos mais complexos e custosos, que podem ser problemáticos quando se trata de grandes bases de dados - devido ao tempo de processamento elevado para calcular as distâncias entre pontos a cada passo da clusterização. Por outro lado, não é exigida a definição *a priori* do número de segmentos desejado, o que pode ser vantajoso nos casos em que não se prevê um número ideal. Entre os métodos hierárquicos mencionados, aglomerativo e divisivo, o primeiro acaba sendo mais utilizado na prática (TAN; STEINBACH; KUMAR, 2006). Isso se deve ao fato de o método divisivo gerar um volume ainda mais elevado de dados, o que dificulta sua aplicação a bases muito grandes.

Os modelos particionais, por sua vez, seguem um processo muito mais simples e podem ser aplicados com softwares mais básicos em comparação aos outros. A complexidade menor permite também que se trabalhe com bases de dados mais volumosas. No entanto, esses métodos exigem o input inicial de um número de clusters desejado, dependem apenas de valores numéricos (pois a partição é feita a partir das médias dos valores) e são igualmente suscetíveis a outliers, uma vez que todos os pontos são designados a algum cluster – mesmo que sejam distantes do centro de todos eles.

### **2.3.1 Métodos hierárquicos aglomerativos**

Para exemplificar os métodos hierárquicos, será apresentado o algoritmo do aglomerativo que é mais popularmente aplicado. Conforme anteriormente explicado, os métodos aglomerativos são iniciados de forma que cada ponto é um cluster, os quais devem ser sucessivamente unidos ao cluster mais próximo até que reste um único cluster (TAN; STEINBACH; KUMAR, 2006)

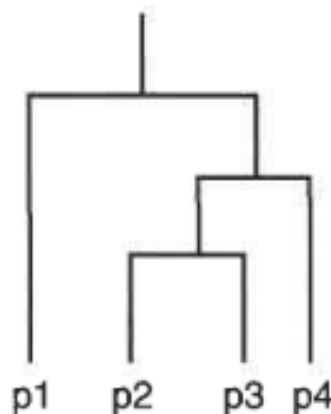
Figura 4: algoritmo de clusterização hierárquica aglomerativa



Fonte: adaptado de Tan, Steinbach e Kumar (2006)

A clusterização hierárquica pode ser representada graficamente por dendogramas, que são diagramas semelhantes a árvores de decisão. Eles mostram os clusters iniciais e a forma como foram posteriormente agrupados (método aglomerativo) ou repartidos (método divisivo), conforme exemplo da figura 5.

Figura 5: exemplo de dendograma



Fonte: extraído de Tan, Steinbach e Kumar (2006)

A principal dificuldade no algoritmo aglomerativo está no cálculo da proximidade entre clusters, que pode ser feita de diferentes maneiras a depender do método escolhido. As cinco principais abordagens para definir-se a distância entre dois clusters, apresentadas por Tan, Steinbach e Kumar (2006), são:

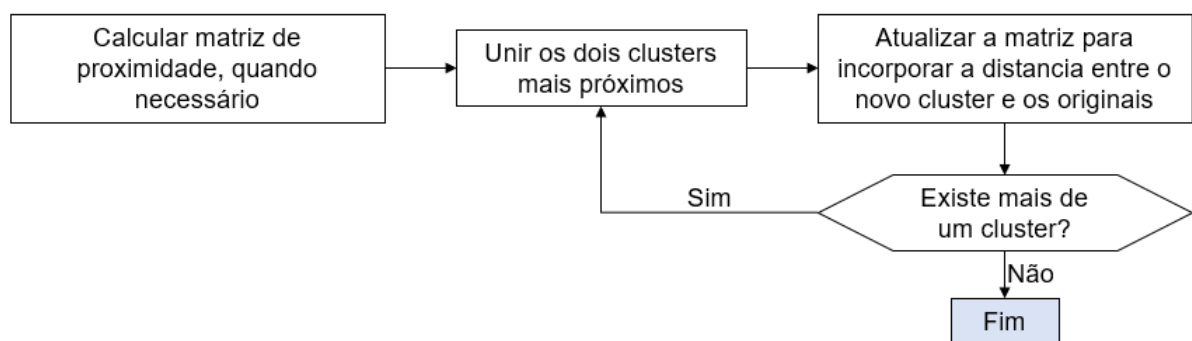
- *Single linkage*: a distância entre os dois pontos mais próximos que estão em clusters diferentes;

- *Complete linkage*: a distância entre os dois pontos mais distantes de clusters diferentes;
- *Average linkage*: a média de todas as distâncias de pares de pontos entre os dois clusters;
- *Centroid linkage*: a distância entre os centroides de cada cluster;
- Método de Ward: ainda com base nos centroides, a distância é calculada pelo aumento da soma dos erros quadrados (SSE) resultante da união dos clusters. A menor distância, nesse caso, é o SSE mínimo.

### 2.3.2 Método de *K-means*

O método de *K-means*, particional, segue uma lógica diferente da hierárquica. A partir de um número  $K$  de clusters, definido pelo usuário, os pontos são designados ao cluster mais próximo até que o centroide dos grupos construídos deixe de variar. O algoritmo dessa técnica é representado no diagrama da figura 6.

Figura 6: algoritmo de *K-means*



Fonte: adaptado de Tan, Steinbach e Kumar (2006)

O algoritmo apresentado costuma ter uma função objetivo que depende das distâncias entre os pontos ou centroides dos cluster. A soma média de quadrados (SSE) pode ser utilizada como medida da qualidade da clusterização, já que quando a distância entre os pontos e o centroide do cluster em que se encontram é a mínima possível, isso significa que os clusters obtidos são a melhor representação dos dados.

Dessa forma, a função objetivo da clusterização pode ser de minimizar a SSE, descrita pela seguinte equação:

$$SSE = \sum_{i=1}^K \sum_{x \in C_i} dist(c_i, x)^2$$

Onde  $K$  é o número de clusters definido,  $x$  é um ponto do cluster  $i$ ,  $c_i$  é o centroide desse cluster e  $dist(c_i, x)$  representa a distância entre o centroide e o ponto. O centroide  $c_i$  de um cluster  $i$  com  $m$  pontos e a distância euclidiana para um espaço com dimensão  $n$  podem ser calculados segundo as equações abaixo, respectivamente.

$$c_i = \frac{1}{m_i} \sum_{x \in C_i} x$$

$$dist(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}$$

Um passo importante para obtenção de um resultado ótimo na clusterização por *K-means* é a escolha de centroides iniciais adequados, pois a escolha aleatória pode gerar grupos subótimos que não necessariamente serão corrigidos pelas iterações futuras. Tan, Steinbach e Kumar (2006) propõem duas alternativas relativamente simples para realizar essa inicialização:

- Realizar uma clusterização hierárquica dos dados e extrair  $K$  clusters, conforme definido pelo usuário. Os centroides desses clusters serão utilizados para inicializar o processo de *K-means*;
- Selecionar um ponto como centroide do primeiro cluster, aleatoriamente, e escolher o ponto mais distante dele como centroide do próximo cluster, repetindo o processo até que se obtenham  $K$  centroides.

O primeiro método é restrito a amostras pequenas e com  $K$  relativamente baixo em comparação ao tamanho da amostra, devido à complexidade e custos de aplicação do método hierárquico. O segundo pode estar suscetível a outliers, problema evitado aplicando-se o método a uma amostra aleatória de dados – que dificilmente apresentará algum outlier devido a sua baixa frequência.



Dolničar (2003) destaca que a literatura é pouco consistente quanto ao número adequado de variáveis a serem usadas nesse método, tendo apenas um autor que apresentou uma restrição mais explícita: segundo Anton Formann (1984, apud DOLNIČAR, 2003), a amostra mínima de dados deve ser igual a  $2^v$ , onde  $v$  é o número de variáveis utilizadas. O ideal, entretanto, é que a amostra seja maior ou igual a  $5 \times 2^v$ .

## 2.4 Validação da clusterização

A avaliação e validação dos clusters obtidos ao final de um processo de clusterização é essencial para garantir que os resultados fazem sentido para o caso estudado. Uma motivação para a realização desse passo é que praticamente todo algoritmo é capaz de dividir os dados em clusters, mesmo que a base de dados não tenha nenhuma estrutura natural de clusterização (TAN; STEINBACH; KUMAR, 2006). Portanto, é importante avaliar a partir de métricas e também de forma qualitativa o resultado da metodologia aplicada.

Existem diferentes critérios que podem ser usados para validar a clusterização, como coesão e separação dos clusters ou medidas combinadas de ambos. Eles avaliam a coesão interna de cada cluster e a qualidade do isolamento entre eles, por meio de diferentes métricas. Apesar de poderem ser medidas de forma independente, Rousseeuw (1987) desenvolveu o conceito de silhuetas que é capaz de medir a coesão e separação dos clusters de maneira conjunta.

Os elementos envolvidos no cálculo da silhueta estão representados na figura 7 e serão apresentados a seguir, considerando um ponto  $i$  que pertence ao cluster A:

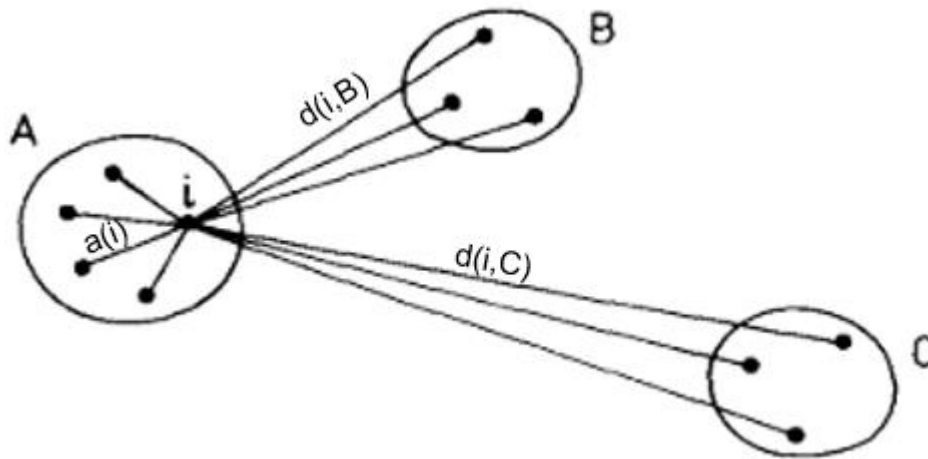
- $a(i)$  = distância média de  $i$  aos outros pontos do cluster A;
- $d(i, C)$  = distância média de  $i$  aos pontos do cluster C, que não contem  $i$ ;
- $b(i) = \min d(i, C)$ , onde C é um cluster diferente de A.

No caso dos clusters representados na figura 7, a distância mínima é obtida entre  $i$  e o cluster B (ou seja,  $b(i) = d(i, B)$ ), que pode ser chamado de vizinho do objeto  $i$ . A silhueta  $s(i)$  é calculada a partir da combinação de  $a(i)$  e  $b(i)$ , de forma que:

$$s(i) = \begin{cases} 1 - a(i)/b(i), & a(i) < b(i) \\ 0, & a(i) = b(i) \\ b(i)/a(i) - 1, & a(i) > b(i) \end{cases}$$

De forma que, para cada objeto  $i$ , pode-se afirmar que  $-1 \leq s(i) \leq 1$ .

Figura 7: elementos envolvidos no cálculo de  $s(i)$ ,  $i \in A$



Fonte: adaptado de Rousseeuw (1987)

É mais fácil entender o significado da silhueta quando se analisam situações extremas. Um  $s(i)$  elevado, próximo a 1, implica que a distância  $a(i)$  é muito menor que  $b(i)$ ; isso evidencia que a clusterização foi bem feita, já que o ponto  $i$  foi designado ao cluster adequado e ainda bastante distante do cluster vizinho. Ao contrário, quando  $s(i)$  aproxima-se de zero, não fica claro se o ponto  $i$  foi designado ao cluster certo ou poderia ter sido colocado no vizinho, sendo um caso intermediário. Por fim, quando  $s(i)$  está próximo de -1, apresenta-se o pior caso possível: significa que o ponto  $i$  está mais próximo dos pontos do cluster vizinho do que dos demais pontos de seu próprio cluster, de forma que não se pode negar que tenha sido mal classificado.

Denomina-se coeficiente de silhueta (SC) a média das silhuetas dos  $n$  pontos analisados, ou seja, a média de  $s(i)$  para  $i=1,2,3,\dots,n$ . De acordo com Kaufman e Rousseeuw (1990), observações levaram a uma interpretação subjetiva do SC para julgar a qualidade da clusterização, apresentada na tabela 1.

Tabela 1: interpretação sugerida dos coeficientes de silhueta

SC	Interpretação sugerida
0,71-1,00	Estrutura robusta
0,51-0,70	Estrutura razoável
0,26-0,50	Estrutura fraca, possivelmente artificial; testar outros métodos
$\leq 0,25$	Nenhuma estrutura substancial

Fonte: adaptado de Kauffman e Rousseeuw (1990)

## 2.5 Sortimento de produto

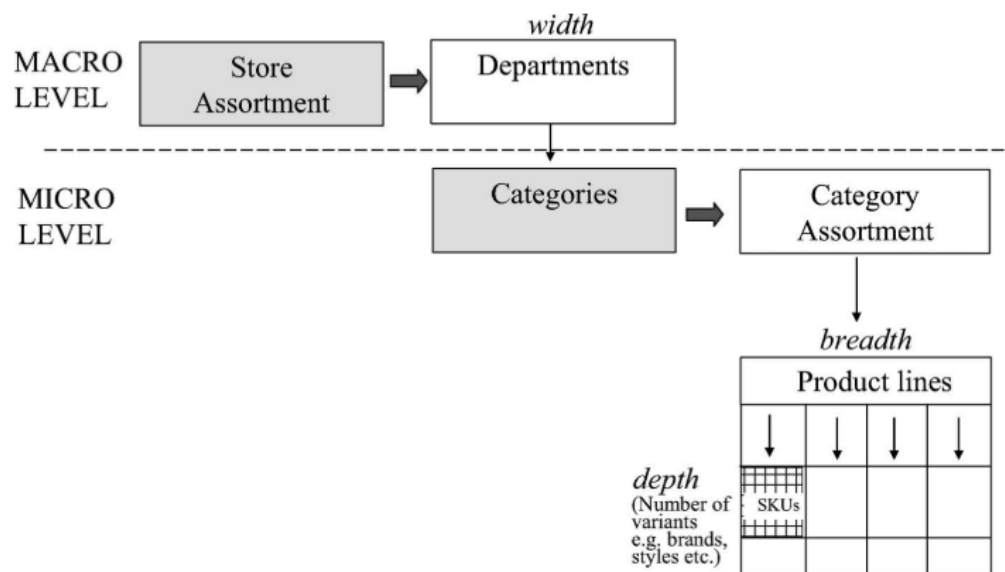
As definições de sortimento de produtos e serviços oferecido por uma empresa são diversas na literatura. Parte delas refere-se ao sortimento de forma mais geral, definido como nada mais que a oferta total de produtos de uma empresa ou loja. Na década de 50, Balderston (1956) definiu sortimento como simplesmente o número total de itens que podem ser vendidos por uma firma em uma transação. Posteriormente, pela definição de Hollander (1996), o sortimento era visto em nível agregado e dizia respeito ao número de linhas de produto ou serviço oferecidas. Em artigos mais recentes, inclui-se a dimensão de profundidade e variedade do sortimento (SIMONSON, 1999), além de características dos produtos, que podem ser independentes, substitutos ou complementares entre si. Todas essas definições referem-se ao sortimento de loja (ou *store assortment*) que define a oferta geral de produtos disponíveis, independentemente de categoria, marca ou outros agrupamentos. Uma outra visão do conceito se relaciona à dimensão de opções disponíveis ao consumidor, na medida em que se associa à gama de produtos oferecidos em suas diferentes variações: cor, tamanho, estilo, sabor (LEE; KUNZ, 2001) - a depender do tipo de produto que se trata. Dessa forma, é introduzida a análise do sortimento em uma visão menos geral, passando do nível loja para categoria ou marca dos produtos estudados.

As definições, no entanto, não são excludentes. É esperado que, para definir o portfólio de produtos completo de uma loja, analise-se o nível mais macro de sortimento por loja que define os departamentos ou categorias a serem comercializados para, posteriormente, especificar os produtos e variações dentro de cada categoria, linha ou marca disponível. São quatro as dimensões essenciais para análise e definição do sortimento de uma loja, segundo o framework de Hart e Rafiq (2006):

- Abrangência: número de diferentes departamentos ou categorias de produtos oferecidos pela loja – a depender da natureza da mesma;
- Variedade: quantidade de diferentes linhas ofertadas dentro de cada departamento ou categoria;
- Profundidade: número de itens disponível em cada uma das linhas, especificando-se tamanho, cor e outras características particulares que resultem na enumeração individual dos SKUs;
- Consistência: refere-se à coesão das categorias e produtos disponíveis do ponto de vista do consumidor.

Essa última dimensão não é diretamente mensurável em número de linhas, categorias ou produtos, sendo mais qualitativa. A consistência do sortimento deve ser analisada pelos funcionários da área de Merchandising da empresa, que terão a sensibilidade necessária para julgar a relação entre as diferentes linhas e categorias de produtos disponíveis. Por ser uma dimensão que não se refere a apenas uma das características do sortimento (categorias, linhas ou marcas e SKUs), a consistência não é enumerada na hierarquia de decisões explicitada na figura 8, diferentemente das demais.

Figura 8: hierarquia de decisões de sortimento



Fonte: extraído de Hart e Rafiq (2006).

O nível estratégico de tomada de decisão, ou nível “macro” na definição de Hart e Rafiq (2006), refere-se à abrangência do sortimento, ou seja, o mix de departamentos ou categorias – a depender do tamanho da loja - que estarão disponíveis. Uma vez definido isso, inicia-se a tomada de decisão do nível “micro” que engloba a determinação da variedade - quais linhas (ou marcas) serão trazidas para a loja - e, finalmente, a profundidade de cada linha, leia-se os itens específicos de cada uma delas a serem vendidos. É importante destacar que a escolha dos itens é o nível mais granular da definição do sortimento da loja, elencando não somente os produtos bem como suas especificidades; cor, tamanho, fragrância e demais características que dependem da natureza do negócio estudado.

Vale notar que, ao contrário do senso comum, nem sempre oferecer o maior número de categorias, linhas, marcas e itens é a melhor estratégia para uma varejista. Além de questões relacionadas ao nível percebido de qualidade e posicionamento da marca, o excesso de variedade ou profundidade pode causar confusão nos consumidores menos experientes, que percebem uma complexidade muito grande na quantidade de opções disponíveis (HART; RAFIQ, 2006).

A definição do sortimento de produtos deve estar alinhada às necessidades e padrão de consumo dos clientes de uma empresa. Apesar de a empresa como um todo ditar certos direcionamentos, as diferentes lojas (ou grupos de lojas com perfis semelhantes) podem apresentar clientes com necessidades distintas que devem ser atendidos da melhor maneira possível. Para tanto, o portfólio de produtos da empresa deve ser editado e direcionado de acordo com as características das lojas que possui.



### 3 FUNDAMENTAÇÃO TEÓRICA SOBRE PREVISÃO DE DEMANDA

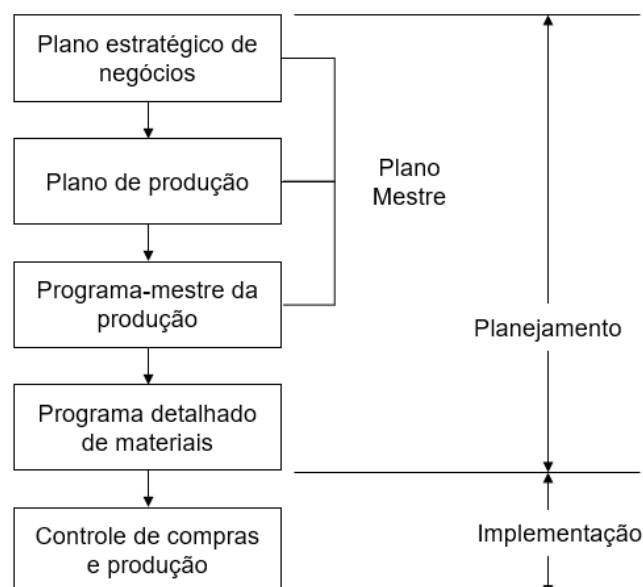
Para embasamento do trabalho realizado, foi feito um levantamento de estudos relacionados a planejamento e previsão, com destaque para referências dos principais autores da literatura.

#### 3.1 Sistema de Planejamento e Controle da Produção (PCP)

Segundo Arnold, Chapman e Clive (2008) é importante que uma empresa organize satisfatoriamente os diversos fatores e recursos envolvidos na produção e comercialização de seus produtos para ser rentável. Um sistema de planejamento adequado deve ser capaz de responder questões relacionadas a priorização e capacidade. A primeira é definida pela demanda: quais, quantos e quando os produtos devem ser fabricados de acordo com as necessidades estabelecidas pelos consumidores. A capacidade, por sua vez, diz respeito aos recursos disponíveis da empresa, que devem ser aplicados para satisfazer a demanda.

Tanto no curto quanto no longo prazo, esses dois fatores devem ser equilibrados, de forma a atender as necessidades de cada nível de planejamento. O planejamento e controle da produção (PCP) pode ser dividido em cinco níveis, representados na figura 9. Os níveis superiores têm um horizonte de planejamento maior e objetivos mais gerais, de longo prazo, que se tornam mais específicos e detalhados conforme se desce na estrutura apresentada, chegando a um horizonte de curto prazo.

Figura 9: planejamento e controle da produção



Fonte: adaptado de Arnold, Chapman e Clive (2008) (2008)

### 3.2 Processo de S&OP

A atualização do plano estratégico de uma empresa deve ser realizada, pelo menos, anualmente (ARNOLD; CHAPMAN; CLIVE, 2008). As alterações no mercado, na demanda e nos recursos da empresa, entretanto, ocorrem com intervalos menores e é necessário atualizar e revisar os planos de cada departamento de acordo com tais mudanças. O S&OP (do inglês, *Sales and Operations Planning*) é o processo responsável por essa revisão, que envolve desde os altos executivos até os times de marketing, finanças, operações e produção e culmina no desenvolvimento do plano de produção.

O processo é iniciado pelos departamentos de vendas e marketing, que comparam as demandas realizadas no último período com a prevista, avaliam o potencial do mercado e calculam a demanda futura. O plano de marketing revisto é então comunicado às áreas responsáveis pela produção, engenharia, finanças que ajustam seus planos de forma a acomodar as novas diretrizes. Caso as mudanças necessárias não sejam possíveis, o plano de marketing deve ser ajustado. Assim, Arnold, Chapman e Clive (2008) afirmam que o plano estratégico de negócios é sempre revisado ao longo do ano, coordenando as atividades dos diversos departamentos envolvidos no S&OP.

Entre as vantagens de implementação de S&OP na empresa, encontram-se a coordenação dos planos dos diversos departamentos, a possibilidade de gerenciamento das mudanças e revisão de planos de acordo com alterações de mercado, culminando na criação de planos realistas e praticáveis e um gerenciamento mais eficiente da produção, estoque e vendas.

### 3.3 Importância de prever

A previsão é parte essencial na tomada de decisões comerciais, operacionais, financeiras e gerenciais de uma empresa. Por meio dela, é possível prever quando um acontecimento se dará ou uma necessidade se manifestará, de forma que sejam tomadas as providências cabíveis para cada situação. Esse processo se mostra essencial em tempos em que os gerentes procuram cada vez mais tomar decisões bem informadas, reduzindo a dependência do sucesso da empresa ao acaso.

Segundo Makridakis, Wheelwright e McGee (1998), a qualidade das previsões em uma área da empresa pode afetar toda a organização, já que os diferentes departamentos estão cada vez mais interligados e dependentes uns dos outros. Um erro elevado na previsão de demanda de produtos finais pode levar a uma previsão desregulada de necessidade de insumos,



orçamento e pricing, por exemplo. Algumas das funções dependentes de previsões em uma organização são:

- Agendamento: o uso eficiente de recursos exige o agendamento de atividades relacionadas a produção, transporte, orçamento, recursos humanos. Previsões de demanda de produtos, matérias primas, trabalhadores, investimento são essenciais para que se realize um agendamento eficiente;
- Aquisição de recursos: o *lead time* de aquisição dos diferentes recursos utilizados por uma organização pode variar. É necessário prever as necessidades com a devida antecedência, de forma que os recursos sejam obtidos a tempo de sua aplicação;
- Determinação de recursos futuros: a decisão de quais e quantos serão os recursos utilizados no longo prazo é muito importante para o planejamento estratégico de uma organização. Fatores como desenvolvimento interno financeiro e tecnológico devem ser previstos para que decisões bem informadas quanto aos recursos futuros sejam tomadas pela gerência.

As três funções acima descritas são exemplos de necessidades de previsão de curto, médio e longo prazo, respectivamente. Por conta dessa diversidade de usos e aplicações das previsões, é necessário que as empresas desenvolvam diferentes modelos de previsão, que sejam interligados e ajustados entre os departamentos para garantir uma coesão nas decisões e ações tomadas.

A eficácia da previsão pode ser questionada devido à dificuldade de se determinar acontecimentos de um futuro incerto. No entanto, os métodos de previsão evoluíram substancialmente ao longo dos anos de forma que diversos eventos são agora facilmente previstos com boa precisão (MAKRIDAKIS; WHEELWRIGHT; MCGEE, 1998). Mais à frente nesse capítulo serão apresentados alguns desses métodos aperfeiçoados e amplamente utilizados pela literatura relevante.

### **3.4 Previsão e Planejamento**

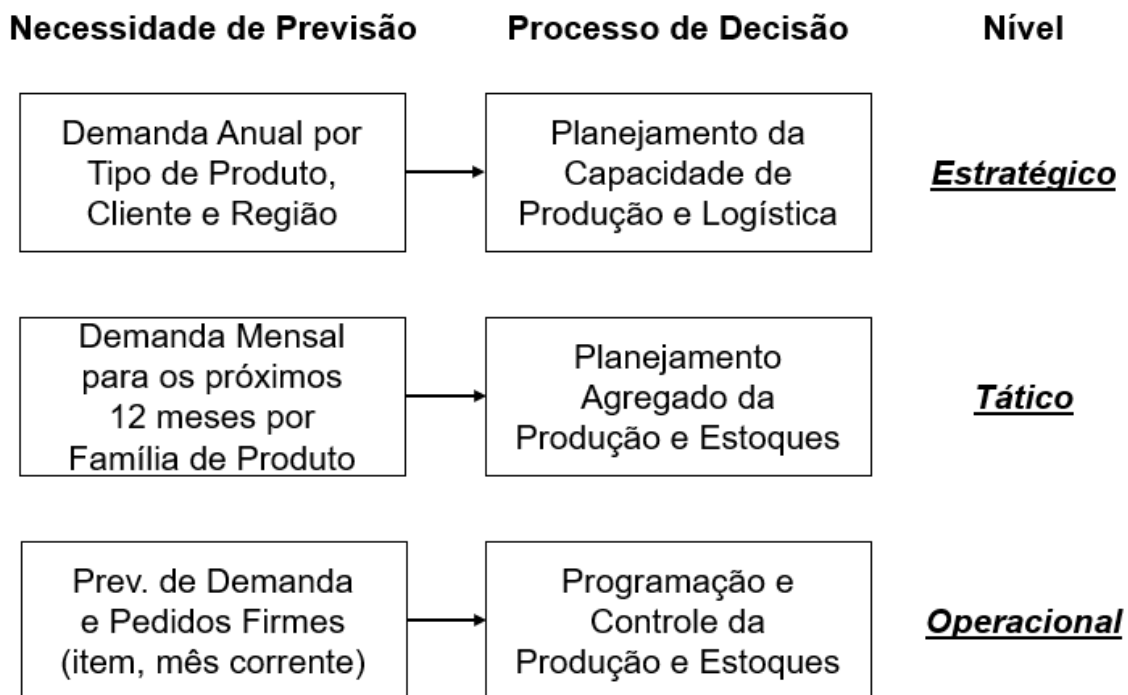
Segundo Mesquita et al. (2008), é com base na previsão da demanda que são tomadas as principais decisões financeiras, comerciais e operacionais de uma empresa. Para cada uma

das decisões é necessária uma visão diferente da previsão, havendo alguns níveis de detalhe e horizontes de previsão possíveis.

Na figura 10 são apresentadas possíveis necessidades de previsão para empresas de manufatura. O planejamento da capacidade de produção e logística exige um nível de detalhe baixo e horizonte de tempo mais alto, por tratar-se de uma decisão estratégica de mais longo prazo. Já o planejamento agregado da produção, no nível tático, exige um nível um pouco mais detalhado (por família de produto, por exemplo) e uma visão de médio prazo. Por fim, no nível operacional, a programação e controle da produção exige que se tenha uma previsão por item do portfólio para um período mais curto e próximo do presente.

É importante ressaltar que as necessidades e horizontes de previsão podem variar conforme o foco da empresa ou área estudada. No ramo do varejo, por exemplo, as decisões estratégicas podem estar relacionadas a expansão das lojas e planejamento logístico enquanto o nível operacional se relaciona a reposição de estoque das lojas.

Figura 10: previsão da demanda no planejamento da produção



Fonte: adaptado de Mesquita et al. (2008)

### 3.5 Princípios de previsão

As previsões devem seguir quatro princípios básicos. Apesar de parecerem óbvios, é essencial entendê-los para realizar um bom uso delas (ARNOLD; CHAPMAN; CLIVE, 2008).

- Previsões normalmente estão erradas. Por tentarem quantificar um futuro incerto, por mais ajustado que seja o método utilizado, a previsão terá um erro inerente. Dessa forma, de nada vale discutir se uma previsão está certa ou errada, mas sim qual opção tem um erro menor.
- Todas as previsões devem incluir uma estimativa de erro. Considerando o princípio anterior, de que as previsões são necessariamente erradas, vale calcular uma estimativa do erro como porcentagem do valor previsto ou um intervalo entre dois valores mínimo e máximo. Essas estimativas podem ser feitas a partir da variabilidade da demanda em relação à média.
- Previsões são mais acuradas para famílias ou grupos. Para Mesquita et al. (2008), previsões de demanda agregada (por categoria, tipo de produto, família etc.) apresentam erros menores do que previsões feitas item a item. Mesmo em grupos com comportamentos bastante constantes, os itens individuais podem se comportar de forma aleatória. Os grupos devem ser divididos de acordo com similaridades entre os itens para o propósito da previsão. No caso de planejamento da produção, por exemplo, a divisão deve ser baseada na similaridade do processo produtivo dos itens (matéria prima, máquinas, técnicas etc.).
- Previsões são mais acuradas para períodos mais próximos. Devido a maior proximidade com o presente, a previsão do futuro próximo é mais simples e próxima do real do que aquela feita para o futuro distante. Isso é de extrema importância para itens com *lead time* longo, sobretudo aqueles que tem demanda instável; a redução do *lead time* melhora a acurácia da previsão.

### **3.6 Procedimento para realização da previsão**

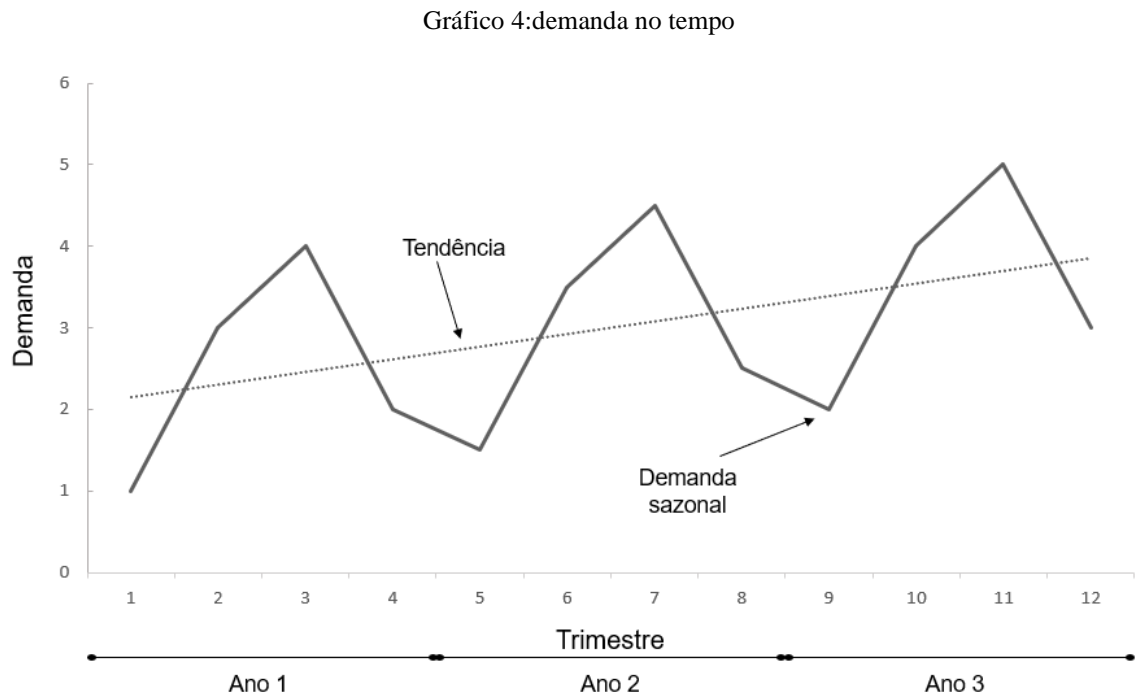
Na literatura sobre o assunto, são propostos alguns procedimentos para a realização de uma previsão com dados quantitativos que seja eficaz, eficiente e acurada. Makridakis, Wheelwright e McGee (1998) propõem o seguinte procedimento:

1. Definição do problema: envolve a compreensão das aplicações da previsão, quem a utilizará e com qual propósito. Antes de estruturar o método e modelo utilizados é essencial entender exatamente qual o objetivo da previsão. Diversas informações são necessárias para essa definição, tal como a amostra de itens a serem previstos, o nível de agregação da previsão, a frequência e período previstos, entre outros;
2. Coleta de informação: trata-se da coleta de bases de dados históricos da variável a ser prevista assim como de conhecimento acumulado pelos experts da empresa. As séries históricas serão utilizadas para confecção do modelo matemático, cujo resultado pode sofrer alterações de acordo com o input dos especialistas;
3. Análise preliminar ou exploratória: etapa na qual se busca entender as características da variável estudada de forma a guiar a escolha do melhor método de previsão. Pode envolver análises de correlação, decomposição, inspeção gráfica etc. Com essas ferramentas deve ser possível identificar a existência de padrões, tendência, sazonalidade, ciclicidade (características da demanda apresentadas no item 3.7 a seguir) e outliers a serem explicados;
4. Escolha e ajuste dos métodos: a análise preliminar direciona a escolha de modelos a partir das conclusões a respeito dos dados, mas as opções devem ser testadas antes de se definir qual será utilizada;
5. Implementação e avaliação do modelo de previsão: a implementação é um passo tão importante quanto a confecção do modelo para as organizações. Não basta que o modelo de previsão seja acurado, a sua utilização deve ser capaz de gerar ações efetivas por parte da gerência da organização. Por exemplo, se a previsão for de redução da demanda, devem ser propostos projetos que impulsionem as vendas para evitar que esse cenário se torne realidade. Além disso, a implementação do modelo inclui a integração dos resultados quantitativos com o julgamento dos especialistas relevantes; na prática, os modelos estatísticos são raramente utilizados sem modificações.

### **3.7 Características da demanda**

Antes de se escolher os métodos de previsão a serem utilizados, é preciso analisar o comportamento da demanda, já que ele afeta profundamente a previsão e a escolha das técnicas

envolvidas. Se analisados os dados históricos graficamente, serão visíveis padrões ou formatos consistentes ao longo do tempo. É evidente que nem todos os pontos da base seguirão os mesmos padrões, mas tendem a aglomerar-se em torno do valor esperado. Existem quatro fatores que definem e explicam os padrões de uma série histórica de demanda como a do gráfico 4: tendência, sazonalidade, ciclos e variação aleatória.



Fonte: adaptado de Arnold, Chapman e Clive (2008)

- **Tendência:** padrão de aumento ou diminuição da demanda entre períodos analisados. No exemplo hipotético acima, percebe-se uma tendência e crescimento – ilustrada pela linha tracejada;
- **Sazonalidade:** responsável pela alteração da demanda com base na época que se analisa. Essa flutuação pode ser influenciada por diversos fatores como datas comemorativas, estação ou outros acontecimentos relevantes periódicos, a depender do tipo de produto estudado. A sazonalidade de um período pode ser percebida pelo cálculo do índice de sazonalidade dele, que demonstra o quando a demanda naquele período está acima ou abaixo da média. É importante notar que a demanda média deve ser calculada com base nos períodos que compõem o comprimento da sazonalidade. Por exemplo, as vendas de equipamentos de ski são mais altas no inverno; considera-se, portanto, que o comprimento da sazonalidade do produto é anual e a demanda média (ou

dessazonalizada) deve ser calculada utilizando-se as vendas nos doze meses (períodos) de um ano:

$$\text{Índice sazonal} = \frac{\text{demanda do período}}{\text{demanda média de todos os períodos}}$$

- Aleatoriedade: variações imprevisíveis ocorrem quando diferentes fatores afetam a demanda durante períodos específicos, que acontecem aleatoriamente;
- Ciclicidade: no longo prazo, flutuações econômicas e sociais podem afetar o padrão de demanda de forma cíclica. O estudo desses padrões, entretanto, se relaciona a previsões mais gerais feitas por agentes econômicos e não ao planejamento realizado pelo departamento comercial de uma empresa.

Além dos fatores enumerados, Arnold, Chapman e Clive (2008) ressaltam que o padrão de demanda de um negócio pode ser estável ou dinâmico na medida em que mantem o mesmo formato ou não ao longo dos períodos. No segundo caso, mudanças imprevisíveis alteram a tendência, sazonalidade ou aleatoriedade da demanda, dificultando a previsão futura.

Além disso, a previsão pode ser dependente ou independente, na medida em que pode ou não ser afetada pela demanda de outros produtos e serviços. Os itens independentes costumam ser produtos finais, cujas partes e componentes são dependentes deles. Nesse caso, é necessário prever somente a demanda dos itens independentes, já que esta pode ser usada como input para o cálculo da demanda das partes que o compõem.

### 3.7.1 Inspeção gráfica

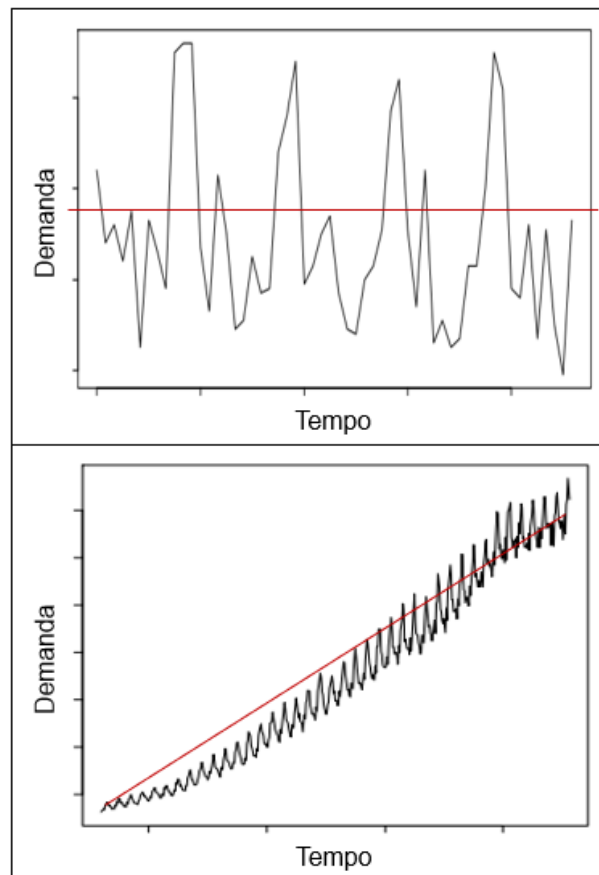
A identificação das tais características da demanda enumeradas anteriormente é essencial para a escolha do modelo adequado de previsão. Na etapa de análise exploratória dos dados, sugerida por Makridakis, Wheelwright e McGee (1998), uma das ferramentas mais poderosas a serem utilizadas é a inspeção gráfica.

Para identificação da tendência de uma série, por exemplo, um método simples e eficaz amplamente utilizado é de analisar o gráfico de demanda por tempo. A simples observação da série permite identificar se a demanda aumenta ou decresce com o passar do tempo, sobretudo quando traçada a linha de tendência linear da série, dada pela equação:

$$Y_t = at + b$$

O gráfico 5 apresenta duas séries históricas teóricas com suas respectivas tendências lineares, em vermelho. Enquanto o gráfico superior não apresenta clara tendência, pode-se identificar com facilidade o crescimento da demanda com o tempo para a segunda série, apresentada no gráfico inferior.

Gráfico 5: gráfico de demanda por tempo para duas séries

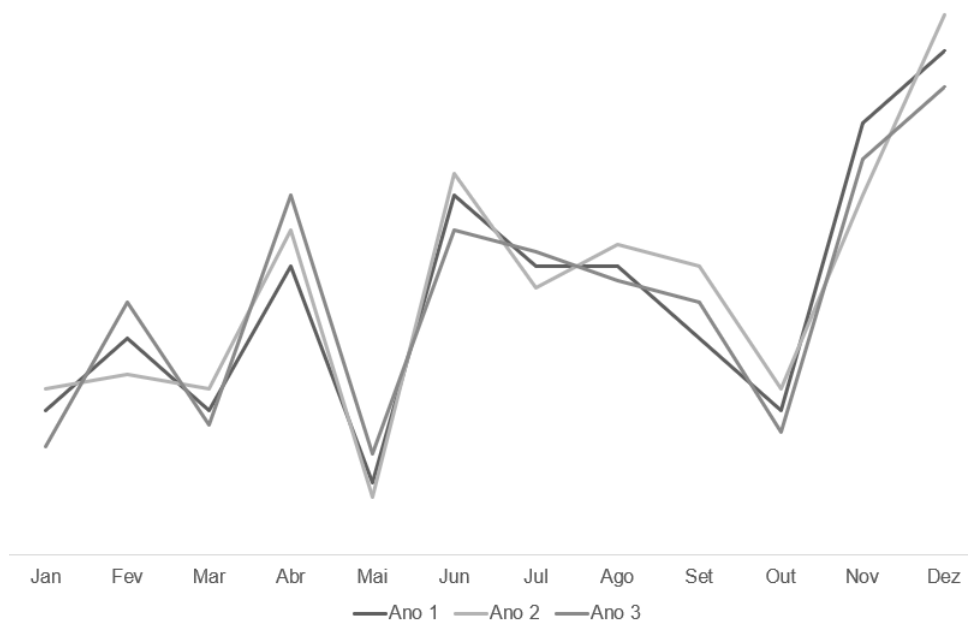


Fonte: elaborado pela autora a partir de dados de Makridakis, Wheelwright e McGee (1998)

A identificação da sazonalidade de uma série pode também ser realizada pela simples análise dos dados históricos, mas existem outras ferramentas razoavelmente simples que contribuem para esse objetivo, como os gráficos de sazonalidade e autocorrelação. Os gráficos de sazonalidade podem ser construídos sobrepondo os dados da série em períodos determinados. A identificação de uma sazonalidade semanal, por exemplo, pode ser realizada sobrepondo-se os dados de demanda diários para diversas semanas. Dessa forma, ficam evidentes os períodos dentro da semana em que a demanda aumenta ou diminui, repetidamente.

O gráfico 6 apresenta o gráfico de sazonalidade de uma série teórica na qual se percebe uma sazonalidade anual. Ao mostrar graficamente as demandas mensais para três diferentes anos é possível identificar claramente, apesar das diferenças entre anos, que existe uma sazonalidade anual na qual a demanda nos meses de novembro e dezembro se eleva de maneira consistente.

Gráfico 6: gráfico de sazonalidade teórico



Fonte: elaborado pela autora.

Já as correlações são medidas capazes de mensurar a relação entre duas variáveis, podendo ser utilizadas para encontrar relações explicativas (MAKRIDAKIS; WHEELWRIGHT; MCGEE, 1998). A autocorrelação tem a mesma função, porém para uma única série de dados.

Pode-se comparar os dados no período  $t$  e  $t-1$ , por exemplo, à procura de identificar a relação entre um valor e seu antecessor direto. Da mesma forma, a comparação pode ser feita entre períodos mais distantes, de acordo com a fórmula abaixo:

$$r_k = \frac{\sum_{t=k+1}^n (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2}$$

Onde:

- $Y_t$  é a demanda no período  $t$

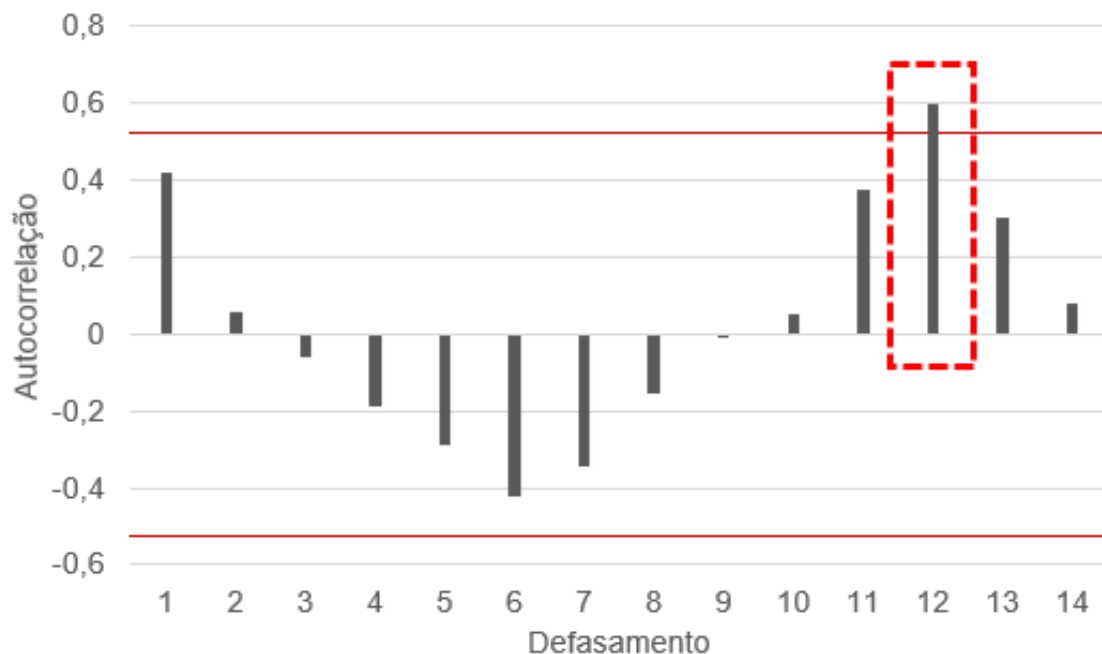


- $\bar{Y}$  é a demanda média da série
- $k$  é a defasagem da autocorrelação

O coeficiente  $r_k$  indica a relação dos valores defasados  $k$  períodos entre si dentro da série analisada. A função autocorrelação é obtida pela união dos coeficientes calculados com diferentes defasagens, que podem ser dispostos em um gráfico de barras chamado correlograma, no qual o eixo horizontal é a defasagem ( $k$ ) e o vertical o coeficiente de autocorrelação correspondente.

A análise do correlograma pode ajudar a identificar padrões nos dados analisados, no que diz respeito, por exemplo, à sazonalidade da série. Suponha que os dados da série representada no gráfico 7 sejam mensais. O valor de autocorrelação é calculado para defasagens de 1 a 14 meses e, pelo correlograma, é perceptível que o coeficiente  $r_{12}$  é o mais elevado entre todos os calculados, indicando uma forte relação entre dados defasados 12 meses, ou seja, uma sazonalidade anual.

Gráfico 7: correlograma de uma série histórica com sazonalidade anual



Fonte: adaptado de Makridakis, Wheelwright e McGee (1998)

Para Makridakis, Wheelwright e McGee (1998) a significância de coeficientes de correlação de uma série com  $n$  observações pode ser estimada pela fórmula  $\pm 2/\sqrt{n}$ , limites apresentados em vermelho no gráfico acima. Assim, os coeficientes  $r_k$  que estejam entre esses

limites podem ser desconsiderados, enquanto os superiores ou inferiores apontam para a existência de padrões na série.

### 3.8 Técnicas de previsão

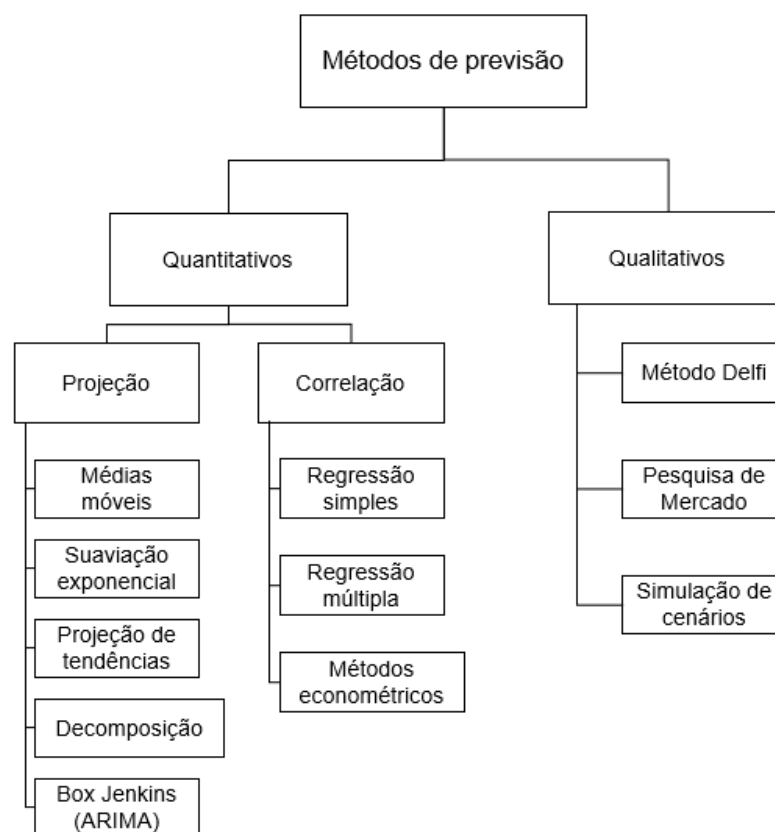
São diversas as possíveis técnicas ou métodos utilizados para a previsão de demanda, mas todos podem ser categorizados entre os seguintes tipos: qualitativos ou quantitativos, que podem ser de projeção ou correlação. A figura 11 apresenta alguns métodos de previsão de demanda comuns em cada uma dessas categorias.

- Métodos qualitativos: são baseados em intuição, julgamento, conhecimento prévio do setor. Por natureza, são subjetivos (ARNOLD; CHAPMAN; CLIVE, 2008). Podem ser utilizados para um nível mais agregado de previsão – de categorias ou famílias de produtos – em um horizonte de tempo mais extenso. Portanto, costumam ser mais utilizados por altos executivos, que tem amplo conhecimento do negócio, para tomada de decisões estratégicas. No caso de previsão da demanda de produtos novos, podem ser usadas pesquisas de mercado e analogias históricas relacionadas a outros produtos similares, por exemplo;
- Métodos quantitativos: são aqueles que tem como base dados quantitativos e técnicas estatísticas para cálculo da demanda prevista. Podem ser de dois tipos, a depender da natureza dos dados utilizados:
  - Métodos de correlação: projeções baseadas em dados externos ao produto, mas relacionados à sua demanda. Possíveis exemplos de dados usados são indicadores socioeconômicos como taxa de nascimentos, consumo de combustível ou taxa de construção civil residencial. Em teoria, a demanda de alguns produtos seria diretamente proporcional ou fortemente correlacionada a alguns desses indicadores; o maior desafio é encontrar um fator cujo comportamento atual explique a demanda futura do item a ser previsto, de forma que não se baseie uma previsão em outra previsão – o que aumenta consideravelmente o erro envolvido. Esses métodos demandam um grande volume de dados históricos de demanda e das variáveis envolvidas na correlação, de forma que se garanta a existência de um padrão de comportamento com base

na relação entre eles. Além disso, são técnicas bastante complexas em termos de modelagem e funcionamento;

- Métodos de projeção de séries temporais: utilizam dados históricos de vendas para realização da previsão. São baseados na suposição de que o que ocorreu no passado deve se repetir. Esses métodos são extremamente importantes e amplamente utilizados para confecção do plano mestre de produção.

Figura 11: exemplos de métodos de previsão de demanda



Fonte: adaptado de Mesquita et al. (2008)

### 3.8.1 Métodos quantitativos de projeção

Em seguida serão apresentados alguns dos principais métodos de projeção para cálculo de previsões, dada a sua relevância no planejamento comercial e operacional das empresas. Nas equações descritas, as variáveis serão:

- $t$  = período a ser previsto
- $p$  = número de períodos utilizados no cálculo da previsão

- P = previsão da demanda
- Y = demanda real

### Média móvel

Esse método utiliza a média da demanda real dos últimos p períodos para previsão do seguinte.

$$P_n = \frac{Y_{t-1} + Y_{t-1} + \dots + Y_{t-p}}{p}$$

A escolha do número p de períodos a ser usado deve ser feita caso a caso, sendo que períodos mais curtos reagem de forma mais rápida às tendências e dão maior peso às informações mais recentes. Por outro lado, quando a demanda varia de forma aleatória e não de acordo com uma tendência, a utilização de um período curto para previsão causará um erro significativo.

Dessa forma, o método de média móvel é mais adequado para prever produtos com demanda estável, sem influência de tendência ou sazonalidade. Para Arnold, Chapman e Clive (2008) médias móveis são úteis para filtrar flutuações aleatórias, quando utilizado o número de períodos adequado.

### Suavização exponencial simples

O método de suavização exponencial utiliza a demanda real assim como a previsão de períodos anteriores para calcular a previsão do período atual. Utiliza-se uma constante de suavização  $\alpha$ , entre 0 e 1, que permite colocar mais peso no período mais recente ou antigo - conforme convier para o caso analisado.

$$P_t = \alpha Y_{t-1} + (1 - \alpha)P_{t-1}$$

A fórmula completa demonstra a forma como os dados históricos perdem relevância conforme sua antiguidade:

$$P_t = \alpha Y_{t-1} + \alpha(1 - \alpha)Y_{t-2} + \alpha(1 - \alpha)^2 Y_{t-3} + \dots$$

A suavização exponencial tem bons resultados para itens com demanda estável, mesmo em horizontes de curto prazo. Nos casos em que há tendência na demanda, ela pode ser melhor identificada e refletida utilizando-se uma constante de suavização mais alta, que confere um peso maior à demanda atual. Ao mesmo tempo, esse valor mais alto pode causar maiores erros caso exista uma flutuação aleatória considerável no padrão de demanda observado. Uma das possíveis soluções é simular diversos fatores  $\alpha$  de forma a identificar aquele que adere melhor aos dados históricos. O maior desafio, quando se trata de qualquer suavização exponencial – tanto a simples quanto as que serão apresentadas a seguir – é encontrar as constantes que minimizem o erro da previsão, aumentando a aderência à demanda real.

#### Suavização exponencial com tendência – Holt's

Proposto por Holt em 1957, esse método é uma variação da suavização exponencial apresentada anteriormente, à qual se agrega um fator capaz de interpretar a tendência intrínseca à série estudada. Dessa forma, além da constante de suavização  $\alpha$ , é introduzida a constante de ponderação para tendência  $\beta$ , também entre 0 e 1. Nesse caso, a equação final da previsão para  $m$  períodos a frente de  $t$  é:

$$P_{t+m} = L_t + b_t \cdot m$$

Onde  $L_t$  é a estimativa do nível de demanda e  $b_t$  a estimativa da tendência no período  $t$ , conforme as equações:

$$\begin{aligned} L_t &= \alpha Y_t + (1 - \alpha)(L_{t-1} + b_{t-1}) \\ b_t &= \beta (L_t - L_{t-1}) + (1 - \beta)b_{t-1} \end{aligned}$$

Inicialmente, é usual utilizar a média das três demandas iniciais como  $b_t$  e a primeira demanda como valor inicial de  $L$  ( $L_0$ ).

#### Suavização exponencial com sazonalidade – Winter's

Essa técnica também é adaptada da suavização exponencial simples, incluindo uma constante de ponderação para sazonalidade  $\gamma$  que varia entre 0 e 1. Introduzido por Winter em 1960 o método, portanto, é adequado para demandas sazonais, que variam conforme a época

dentro do ciclo (podendo o ciclo ser um dia, semana, mês, ano etc.) e não sigam uma tendência de crescimento ou redução. Considerando que a sazonalidade da série pode ser multiplicativa ou aditiva, as equações para implementação dessa metodologia variam.

- Sazonalidade multiplicativa: considerando-se uma sazonalidade de comprimento  $s$ , a previsão para o período  $t+m$  é dada por

$$P_{t+m} = L_t \cdot S_{t-s+m}$$

Onde  $L_t$  é a estimativa do nível de demanda e  $S_t$  a sazonalidade no período  $t$ , conforme as equações:

$$L_t = \alpha \frac{Y_t}{S_{t-s}} + (1 - \alpha)(L_{t-1})$$

$$S_t = \gamma \frac{Y_t}{L_t} + (1 - \gamma)S_{t-s}$$

- Sazonalidade aditiva: considerando-se uma sazonalidade de comprimento  $s$ , a previsão para o período  $t+m$  é dada por

$$P_{t+m} = L_t + S_{t-s+m}$$

Onde  $L_t$  é a estimativa do nível de demanda e  $S_t$  a sazonalidade no período  $t$ , conforme as equações:

$$L_t = \alpha(Y_t - S_{t-s}) + (1 - \alpha)(L_{t-1})$$

$$S_t = \gamma(Y_t - L_t) + (1 - \gamma)S_{t-s}$$

#### Suavização exponencial com tendência e sazonalidade (Holt Winter's)

O método de Holt Winter's é uma espécie de união dos dois métodos anteriormente apresentados, sendo uma suavização exponencial que incorpora constantes de ponderação tanto para tendência ( $\beta$ ) quanto sazonalidade ( $\gamma$ ) da amostra. Portanto, essa técnica trabalha com três equações: de nível ( $L_t$ ), tendência ( $b_t$ ) e sazonalidade ( $S_t$ ).

Existem duas formas de se modelar a sazonalidade para utilização da técnica apresentada, assim como para o método de Winter's. Elas diferem no cálculo dos índices de sazonalidade, que afetam o nível e previsão posteriormente.

- Holt Winters Multiplicativo: considerando-se uma sazonalidade de comprimento  $s$ , a previsão para o período  $t+m$  é dada por

$$P_{t+m} = (L_t + b_t \cdot m)S_{t-s+m}$$

Onde:

$$L_t = \alpha \frac{Y_t}{S_{t-s}} + (1 - \alpha)(L_{t-1} + b_{t-1})$$

$$b_t = \beta(L_t - L_{t-1}) + (1 - \beta)b_{t-1}$$

$$S_t = \gamma \frac{Y_t}{L_t} + (1 - \gamma)S_{t-s}$$

- Holt Winters Aditivo: considerando-se uma sazonalidade de comprimento  $s$ , a previsão para o período  $t+m$  é dada por

$$P_{t+m} = L_t + b_t \cdot m + S_{t-s+m}$$

Onde:

$$L_t = \alpha(Y_t - S_{t-s}) + (1 - \alpha)(L_{t-1} + b_{t-1})$$

$$b_t = \beta(L_t - L_{t-1}) + (1 - \beta)b_{t-1}$$

$$S_t = \gamma(Y_t - L_t) + (1 - \gamma)S_{t-s}$$

Para iniciar a previsão no método de Holt Winters é preciso estimar valores iniciais dos índices utilizados. O nível inicial é calculado usando no mínimo um ciclo completo de sazonalidade ou temporada, ou seja,  $s$  períodos (MAKRIDAKIS; WHEELWRIGHT; MCGEE, 1998).

$$L_s = \frac{1}{s}(Y_1 + Y_2 + \dots + Y_s)$$

Para o índice de tendência, é interessante utilizar pelo menos dois ciclos de sazonalidade completos. Em caso de restrição de dados, podem ser utilizados métodos alternativos para inicialização desse índice.

$$b_s = \frac{1}{s} \left( \frac{Y_{s+1} - Y_1}{s} + \frac{Y_{s+2} - Y_2}{s} + \dots + \frac{Y_{s+s} - Y_s}{s} \right)$$

Os índices de sazonalidade, por fim, são inicializados de maneiras diferentes a depender do tipo de sazonalidade considerada: multiplicativa ou aditiva. Para a primeira, o índice é calculado utilizando a razão das demandas dos primeiros períodos sobre a demanda média da temporada ou nível inicial. No caso da aditiva, utiliza-se a diferença entre esses valores.

- Método multiplicativo:  $S_1 = \frac{Y_1}{L_s}$     $S_2 = \frac{Y_2}{L_s}$    ...    $S_s = \frac{Y_s}{L_s}$
- Método aditivo:  $S_1 = Y_1 - L_s$     $S_2 = Y_2 - L_s$    ...    $S_s = Y_s - L_s$

### 3.8.2 Ajustes e transformações

Para evitar conclusões equivocadas sobre uma série ou previsões enviesadas por conta de erros de dados, é importante realizar alguns ajustes nos dados estudados. Algumas das transformações mais relevantes são as relacionadas ao calendário, já que a extensão de um mês pode variar até 10% - de 28 a 31 dias. A demanda mensal corrigida  $W_t$  pode ser facilmente obtida pela seguinte equação:

$$W_t = Y_t \times \frac{\text{número de dias médio de um mês}}{\text{número de dias do mês } t} = \frac{365,25/12}{\text{número de dias do mês } t}$$

Da mesma forma, no caso de indústrias como varejo ou serviços financeiros, a série deve ser corrigida pelas alterações de dias úteis ou comerciais do mês em questão. Esses ajustes são necessários porque o número de dias de funcionamento do negócio em um mesmo mês pode variar ano a ano. Considerando-se o caso mais simples, em que um dia é classificado simplesmente como comercial ou não, a demanda corrigida  $W_t$  pode ser obtida pela seguinte equação:



$$W_t = Y_t \times \frac{\text{número de dias comerciais médio de um mês}}{\text{número de dias comerciais do mês } t}$$

Onde se considera que a demanda inicial  $Y_t$  já foi corrigida pela extensão do mês, conforme anteriormente apresentado.

### 3.9 Acurácia da previsão

Para escolher o melhor método a ser usado na previsão de demanda, é importante medir os erros resultantes de cada um deles e compará-los em busca daquele que apresente maior acurácia. Serão apresentadas a seguir algumas medidas estatísticas do erro de uma previsão que podem ser usadas com esse fim.

#### Erro

Para cada período  $t$ , o erro pode ser calculado como a diferença entre a demanda real e a previsão

$$e_t = Y_t - P_t$$

Essa equação descreve o erro em um único período  $t$ . Para analisar a acurácia de uma amostra de  $n$  períodos deve ser usada uma das equações abaixo.

#### Erro médio (ME)

O erro médio ou ME (do inglês, *mean error*) nada mais é do que a média dos erros individuais da amostra de períodos  $t$  analisada.

$$ME = \frac{1}{n} \sum_{t=1}^n e_t$$

Ao somar erros individuais positivos e negativos, a medida cancela valores de sinais opostos. A magnitude do erro da previsão, portanto, não é bem calculada pelo erro médio, mas o seu viés sim. Ou seja, um erro médio elevado positivo ou negativo evidencia que o modelo está sistematicamente subvalorizando ou supervalorizando a demanda, respectivamente.

### Erro médio absoluto (MAE)

Diferentemente do erro médio, o MAE (do inglês, *mean average error*) utiliza valores absolutos. Dessa forma, evita a anulação de erros de sinais opostos.

$$MAE = \frac{1}{n} \sum_{t=1}^n |e_t|$$

### Erro quadrático médio (MSE)

O MSE (*mean squared error*) desempenha uma função semelhante á do MAE na medida em que evita a anulação de erros positivos e negativos. Enquanto o MAE utiliza o módulo dos erros individuais, o MSE os eleva à segunda potência. As vantagens do segundo são, sobretudo, a facilidade de compreensão e utilização do método. Esse indicador é amplamente utilizado para comparar erros de uma mesma série entre modelos ou métodos, já que elimina a anulação de valores.

$$MSE = \frac{1}{n} \sum_{t=1}^n e_t^2$$

Todos os métodos até aqui apresentados utilizam medidas dependentes da escala dos dados trabalhados, sendo pouco relevantes na comparação de diferentes períodos ou objetos de estudo. Para esses fins, devem ser utilizadas as medidas relativas, apresentadas a seguir.

### Erro percentual

Similar ao erro individual, mas transformado em percentual.

$$PE_t = \left( \frac{Y_t - P_t}{Y_t} \right) \cdot 100$$

### Erro percentual médio (MPE)

Semelhante ao erro médio, o MPE (*mean percentage error*) pode ser utilizado para determinar o viés da previsão.

$$MPE = \frac{1}{n} \sum_{t=1}^n PE_t$$

### Erro percentual médio absoluto (MAPE)

Semelhante ao erro médio absoluto, o *MAPE* usa os valores absolutos dos erros percentuais individuais em seu cálculo. Por eliminar a anulação de fatores e relativizar o erro, é a medida mais utilizada para comparar métodos de previsão independentemente do tamanho da amostra e período considerado.

$$MAPE = \frac{1}{n} \sum_{t=1}^n |PE_t|$$

É importante notar que a acurácia do modelo apenas pode ser medida a partir de erros de dados que não foram utilizados para parametrização dele. Os erros em relação à série histórica poderiam ser minimizados por meio de um ajuste forçado dos parâmetros aos dados, sendo necessário que parte da série seja isolada e utilizada somente para realização dos testes de acurácia e não para calibragem do modelo. Por conta disso, Makridakis, Wheelwright e McGee (1998) propõem uma divisão da série histórica em dois conjuntos: inicialização e teste ou *holdout*.

O primeiro conjunto é utilizado para estimar valores iniciais de coeficientes (como nível, tendência e sazonalidade no caso de suavizações exponenciais) e calibrar os parâmetros de acordo com o comportamento dos dados históricos, de forma a reduzir os erros entre os valores reais e previstos.

Os dados separados para teste ou *holdout set*, são aqueles a partir dos quais se pode medir a acurácia real do modelo. A previsão é realizada para os períodos correspondentes sem que os dados históricos tenham sido utilizados na calibragem do modelo, sendo seus erros decorrentes de previsões genuínas.

Dessa forma, apenas os erros relacionados ao *holdout set* devem ser utilizados para comparação e avaliação de modelos. Em geral, reserva-se uma porcentagem  $p$  dos dados para calibragem e  $(1-p)$  para teste, sendo  $p > 0,5$ . É comum reservar pelo menos 2/3 dos dados para calibragem do modelo.

### 3.10 Seleção dos métodos de previsão

A seleção de um método de previsão de demanda não deve, de forma alguma, limitar-se àquele com a maior acurácia ou complexidade e sofisticação estatística. Pelo contrário, a escolha depende de quatro fatores relacionados ao objetivo e função da previsão: os dados (projeção ou correlação), as características dos dados, o tipo de dado (comprimento dos períodos) e a frequência da previsão (MAKRIDAKIS; WHEELWRIGHT; MCGEE, 1998).

A escolha entre métodos de projeção, que utilizam somente a série histórica como base, ou correlação que depende de outras variáveis, é subjetiva. Pode ser interessante utilizar modelos de regressão quando a empresa tiver interesse em entender melhor como certas decisões podem impactar a demanda, mas é necessário um volume muito maior de dados para utilização dessa técnica em comparação às projeções temporais. Quando o objetivo da empresa é apenas de obter uma previsão, os modelos de projeção de séries temporais são muito mais baratos e simples de usar sendo, na média, também mais acurados (MAKRIDAKIS; WHEELWRIGHT; MCGEE, 1998).

É essencial, também, que o modelo de previsão utilizado incorpore as características percebidas na série analisada. Ou seja, caso a análise exploratória dos dados evidencie forte fator de sazonalidade ou tendência na série, é importante que sejam escolhidas técnicas que levem em consideração tais características para realização da previsão de forma confiável.

A respeito do tipo de dado, Makridakis, Wheelwright e McGee (1998) afirmam que, em geral, a aleatoriedade costuma diminuir conforme aumenta o nível de agregação de uma série. Portanto, dados anuais são mais consistentes do que os bimestrais que, por sua vez, contém menos aleatoriedade do que os mensais e assim sucessivamente. Quando possível, dependendo do objetivo da previsão, se deve priorizar a utilização de períodos mais espaçados de previsão para reduzir o fator aleatório obtendo, dessa forma, uma série de dados mais consistente para a qual sejam adequados métodos mais complexos e sofisticados de previsão.

Por fim, a frequência das previsões é outro fator importante para a escolha do método mais adequado. Isso se deve ao fato de que, quanto menor a frequência da análise, o número de previsões realizadas em um período determinado é, conseqüentemente, inferior. Dessa forma, os responsáveis pelo processo podem desprender mais tempo e esforço em previsões menos frequentes enquanto aquelas realizadas repetidamente devem ser simples e tomar o menor tempo possível. Técnicas de previsão mais completas, sofisticadas e complexas são indicadas para previsões anuais ou bimestrais, por exemplo, enquanto aquelas mais simples e diretas devem ser utilizadas em casos de previsão diária ou semanal.

Além desses fatores, é de extrema relevância a estrutura e capacidade de tratamento de dados da empresa que realizará a previsão em base regular. A metodologia utilizada deve adaptar-se às necessidades de planejamento ao mesmo tempo em que atende às restrições de software, nível de complexidade e tempo de dedicação do time que será responsável pelo seu desdobramento.



## 4 SEGMENTAÇÃO DE LOJAS E DEFINIÇÃO DO SORTIMENTO DA VAREJO SA

A seguir serão apresentados o processo e os resultados da segmentação das lojas da Varejo SA. A empresa havia realizado uma divisão anterior ao projeto que utilizava demasiados critérios, gerando uma complexidade desnecessária e grupos pouco distintos ou representativos.

Foi realizada uma comparação com outros grandes varejistas que já foram clientes da consultoria Bain & Company, cuja rede de lojas é significativamente maior e mais diversa que a da empresa estudada. Identificou-se que grande parte das empresas, assim como a Varejo SA, utiliza variáveis relacionadas a tamanho e performance das lojas, classe social e clima da região em que se localizam. Entretanto, todas as empresas analisadas utilizam menos variáveis e gerenciam um menor número de segmentos, apesar da maior complexidade de suas redes.

Figura 12: benchmark de segmentação de canais

		Varejista #1 (Moda)	Varejista #2 (Moda)	Varejista #3 (Calçados)	Varejista #4 (Casa)	Varejista #5 (Não-alimentar)	Varejo SA (Acessórios)
	# Segmentos	6	8	8	5	7	18
	# Lojas	~300	~400	~330	~160	~570	<100
Tamanho e Performance	Receita		✓			✓	
	Margem		✓				
	Giro					✓	
	Pequena	✓			✓		✓
	Média				✓		✓
	Grande	✓			✓		✓
Classe social	A			✓	✓		✓
	AB			✓	✓	✓	✓
	BC			✓	✓	✓	✓
Clima	Frio	✓	✓	✓			✓
	Hib. Frio	✓	✓	✓			✓
	Quente	✓	✓	✓			✓
	Hib. Quente		✓	✓			✓

Fonte: elaborado pela autora.

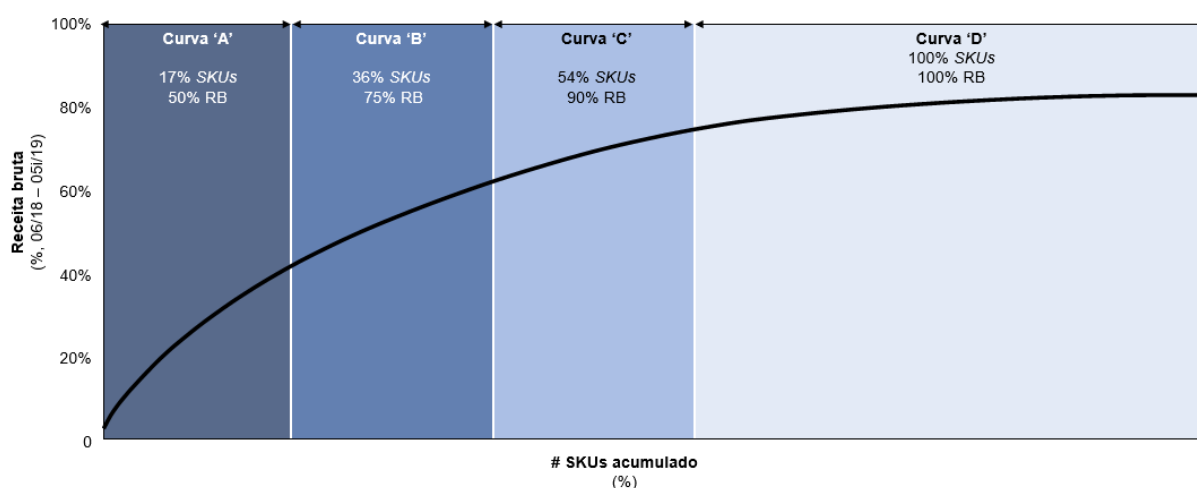
### 4.1 Propósito da segmentação

A Varejo SA, como mencionado no capítulo introdutório do trabalho, atuava em outros canais de venda antes de iniciar a comercialização no varejo físico. Por conta da variedade de clientes e modelo de negócios com que a empresa trabalha, o sortimento de produtos para o

varejo deveria ser adaptado em relação aos demais canais e, inclusive, entre as diversas lojas. É importante garantir que as lojas recebam um sortimento de produtos adequado ao seu tamanho, público, fluxo ou outras características a serem estudadas. Como destacado na revisão bibliográfica, a falta de direcionamento do sortimento oferecido pode levar à confusão dos consumidores e baixo desempenho das lojas.

O envio de todos os SKUs disponíveis para a rede culminou em uma receita altamente concentrada, havendo 54% dos SKUs responsáveis pela quase totalidade das vendas (90%). Essa concentração evidencia que existe a possibilidade de edição de sortimento, de forma a retirar das lojas aqueles produtos cujo volume de venda e receita são muito baixos, os quais geram complexidade desnecessária na gestão comercial do varejo.

Gráfico 8: gráfico acumulado de receita bruta por SKU, Jun/18-Mai/19



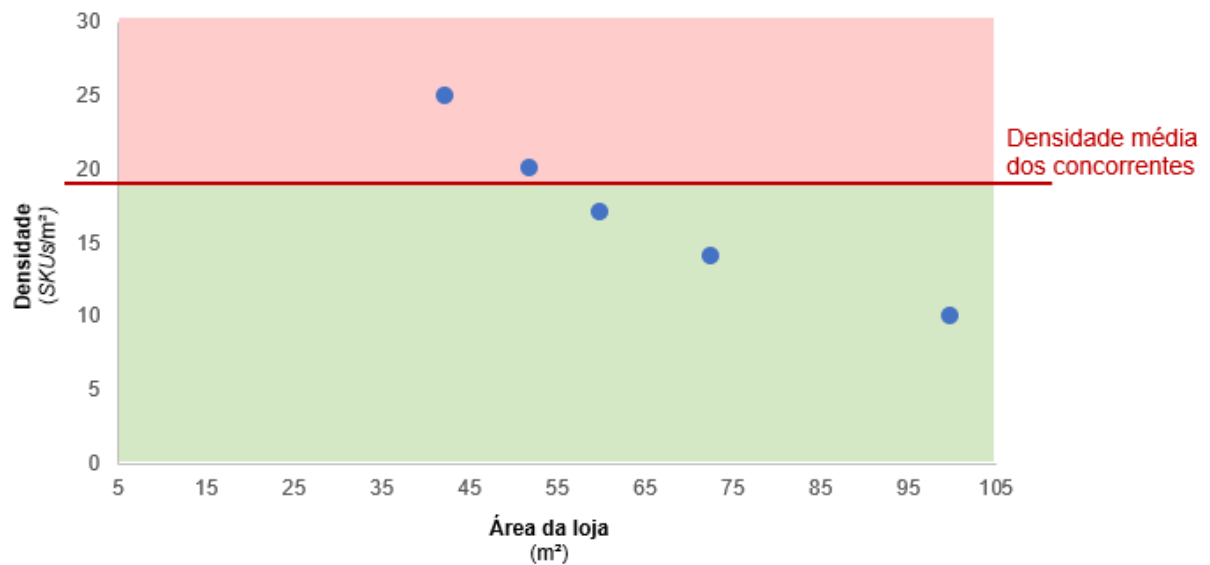
Fonte: elaboração da autora a partir de dados da Varejo SA.

Além da concentração da receita, outro fator importante causado pela falta de edição de sortimento foi o desequilíbrio da densidade de itens expostos entre as lojas. Isso ocorreu pois, independentemente da área das lojas, todas receberam a mesma variedade de produtos que, pelas diretrizes da empresa, devem necessariamente ser expostos para que possam despertar o interesse dos consumidores. Assim, as menores lojas tiveram que comprimir os itens para que todos pudessem ser expostos ao mesmo tempo, culminando em uma densidade elevada (SKUs/m<sup>2</sup>) em comparação às grandes. Essa concentração de produtos cria uma experiência de compra inferior para o consumidor, uma vez que as lojas lotadas parecem menos organizadas e *premium*. No gráfico 9, algumas lojas da rede foram representadas de acordo com a densidade de itens expostos (SKUs/m<sup>2</sup>) e a área da loja, evidenciando o aumento da primeira conforme ocorre a diminuição da área.



O valor de densidade adequado é relativo e depende do tipo de produto e características da marca, mas pode ser comparado ao das concorrentes diretas. Tais valores foram obtidos em campo, por meio de visitas a diversas lojas. A densidade média delas foi traçada no gráfico, de forma a representar um parâmetro para a rede da Varejo SA.

Gráfico 9: gráfico de densidade por área de lojas selecionadas da Varejo SA



Fonte: elaborado pela autora.

Portanto, um dos principais objetivos da segmentação das lojas da Varejo SA é fornecer uma base para edição do sortimento, de forma a reduzir a densidade das lojas e melhorar suas performances, evitando o acúmulo de itens que não são demandados pelos consumidores.

Além disso, a segmentação das lojas deve também funcionar como insumo para o departamento de marketing desenvolver diferentes iniciativas de publicidade nos pontos de venda, desde descontos e anúncios até planos de exposição e treinamentos personalizados, em busca de oferecer uma melhor experiência aos clientes e, assim, impulsionar os lucros da empresa.

## 4.2 Exploração das variáveis a serem consideradas

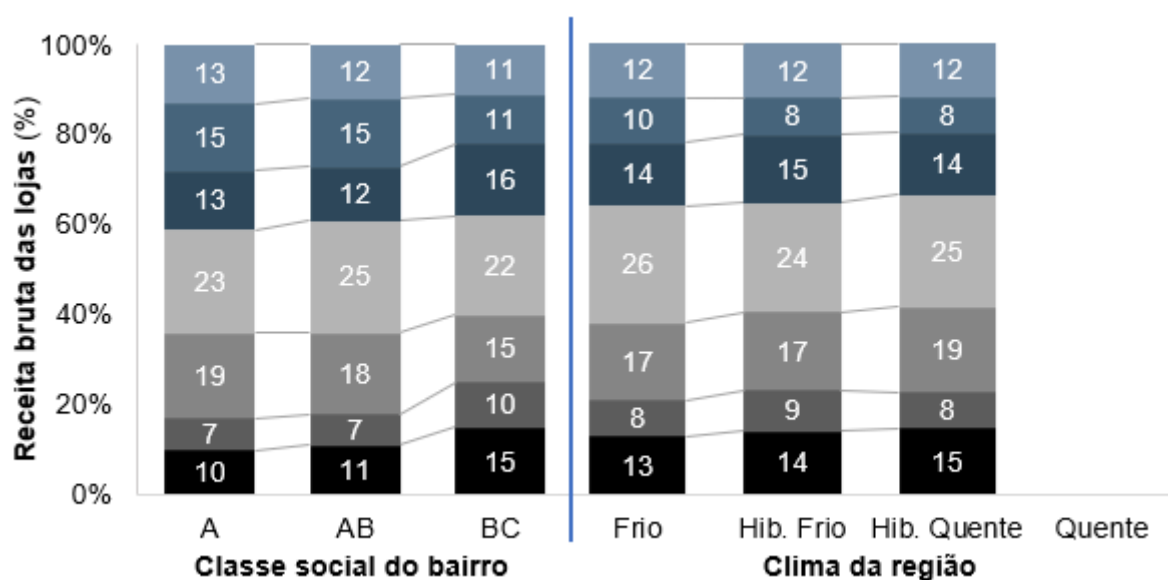
A segmentação das lojas da Varejo SA deve ser simples e atualizável, de forma que o time possa revisar o modelo facilmente conforme ocorram alterações na rede, como abertura ou fechamento de unidades. Para que seja acionável, uma das características principais de uma boa segmentação segundo Blythe (2005), é necessário que os critérios utilizados sejam capazes

de dividir o grupo de acordo com as necessidades e propósito da segmentação. No caso estudado, um dos principais problemas a serem resolvidos é a diferença da densidade de itens expostos, que está intimamente vinculada à área da loja (m<sup>2</sup>); esse fator, portanto, deve ser utilizado para a segmentação. O faturamento anual também pode ser utilizado, principalmente como fator de priorização para tomada de decisões estratégicas e de marketing.

Dada a aplicação principal da segmentação no caso da Varejo SA, que é a edição de sortimento, percebe-se a necessidade de selecionar variáveis que sejam capazes de diferenciar significativamente os segmentos de acordo com seus padrões de consumo. Foram testadas variáveis relacionadas a classe social do bairro e clima da região das lojas em relação ao mix de categorias e linhas de produtos vendidos.

Pelo gráfico da 10, fica evidente que o mix de categorias vendidas varia muito pouco entre as lojas de classe A e AB, com uma alteração um pouco mais relevante em relação às lojas localizadas em bairros de classe BC. As vendas por categoria de lojas A+AB comparadas às BC apresentaram um desvio padrão médio de 2,2%. No que diz respeito ao clima, a variação do mix de categorias é menos relevante, com um desvio médio de 0,7%. Vale notar que ainda não há lojas abertas em regiões de clima quente, mas a tendência é que o seu comportamento comprove a pequena influência do clima, com base no que foi percebido até o momento.

Gráfico 10: gráfico de receita bruta (%) por categoria das lojas da Varejo SA, 2018

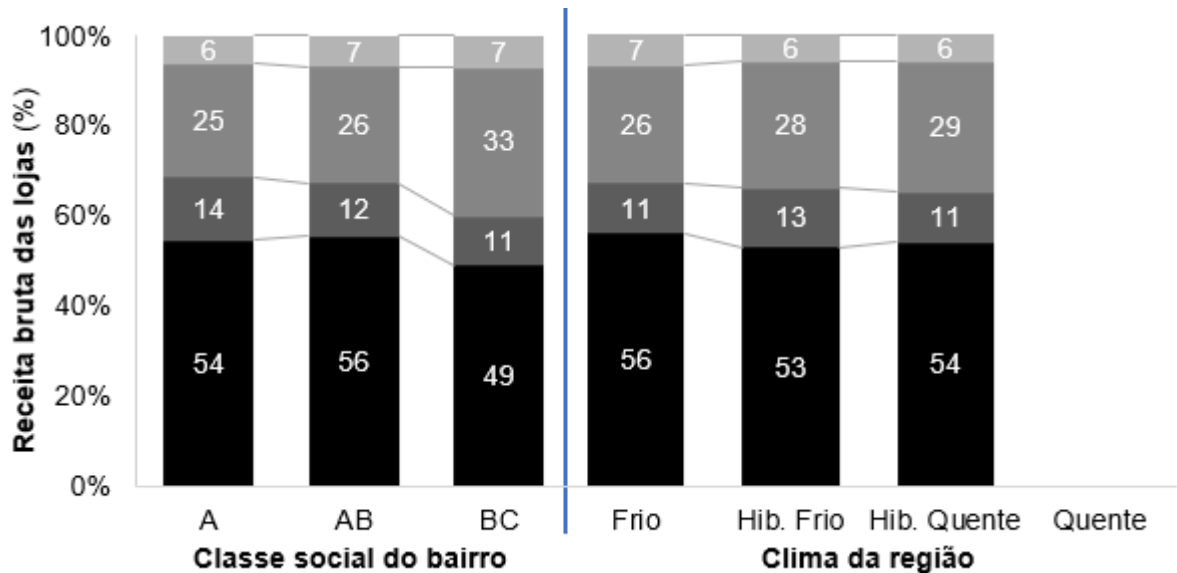


Fonte: elaborado pela autora a partir de dados da Varejo SA

Observando o gráfico 11, o mix de linhas da categoria mais vendida da Varejo SA – aquela que representa mais de 20% da receita em todos os climas e classes – o mesmo

comportamento pode ser identificado: existe uma variação relevante entre o mix vendido em lojas de classe BC comparado às outras (desvio médio de 2,8%) e a diferença de clima tem mínima influência no mix de linhas, com desvio médio de 1,1%.

Gráfico 11: gráfico de receita bruta (%) por linha da principal categoria da Varejo SA



Fonte: elaborado pela autora a partir de dados da Varejo SA

#### 4.3 Escolha do método de clusterização

Atualmente, a Varejo SA tem poucas dezenas de lojas no Brasil e grande parte delas tem características semelhantes: são localizadas na região sudeste, em shoppings em bairros de classe AB, com alto fluxo. Como apresentado na introdução, entretanto, a empresa tem planos de expandir expressivamente a rede pelo país, tanto em número de lojas quanto em variedade de suas características. Portanto, é importante desenhar um processo de segmentação eficiente não apenas para as poucas lojas atuais, mas que possa ser empregado pelos próximos anos e incluir as novas unidades. Apesar de robusto, o processo de segmentação deve ser simples e facilmente atualizável, além de contar com um número de grupos pequeno que mantenha a complexidade de gestão baixa. Outra característica desejável é que seja possível representar os grupos segmentados em gráficos, o que facilita a visualização e materialidade da divisão.

Por conta dessas restrições e da escolha de software da empresa, que insistiu em trabalhar somente com Microsoft Excel, foi selecionado o método de clusterização de *K-means* com as variáveis área (m<sup>2</sup>) e faturamento das lojas (R\$) no último ano. Foram utilizados 18 pontos na segmentação, valor que apesar de não se igualar ao tamanho ideal da amostra quando

utilizadas duas variáveis ( $5 \times 2^2 = 20$ ), supera o mínimo proposto por Formann (1984, apud DOLNIČAR, 2003) que seria igual a 4 ( $2^2 = 4$ ).

A importância da classe social do bairro para a definição do sortimento das lojas foi apresentada, mas a variável será posteriormente considerada de forma que seja levada em conta apenas nos segmentos de menor porte. As lojas maiores têm capacidade de receber todo o portfólio de produtos mantendo uma densidade de exposição adequada, portanto, não existe necessidade de aumentar a complexidade de gestão dos segmentos e os erros do método aplicado agregando essa variável inicialmente.

Uma característica importante de métodos particionais, como o *K-means*, é a necessidade de selecionar o número de clusters  $k$  a serem alcançados no final do processo. Apesar de ser uma restrição, no caso estudado se configura como uma vantagem, já que a necessidade de gerar uma base de fácil gerenciamento limita o número total de segmentos. Dessa forma, decidiu-se testar a divisão da rede de lojas com  $k=3, 4$  ou  $5$ .

As variáveis utilizadas, ambas contínuas, foram transformadas de forma que o dado de maior valor fosse igual a 100 e os demais, proporcionais a ele. Assim, as duas variáveis assumiram ordens de grandeza iguais, evitando que uma delas dominasse o cálculo das distâncias Euclidianas utilizadas pelo método.

A inicialização, assim como proposto por Tan, Steinbach e Kumar (2006) foi feita por um método simples, mas capaz de alcançar bons resultados em uma base pequena e regular como a utilizada: um ponto foi escolhido como centroide do primeiro cluster, aleatoriamente, aquele mais distante dele foi selecionado para o segundo cluster e assim sucessivamente até esgotar-se o número de clusters.

A seguir, serão apresentados os resultados obtidos para os testes realizados com os diversos números de clusters. A seleção da melhor opção será baseada na validação quantitativa, por meio da análise do coeficiente de silhueta, assim como qualitativa. Essa última deve garantir que a divisão obtida cumpra os cinco critérios de uma boa segmentação propostos por Blythe (2005): mensurável, acessível, substancial, coerente e estável.

#### **4.4 Resultados e validação da clusterização**

O método de *K-means* foi testado para um número de clusters igual a 3, 4 e 5. Seguindo o procedimento conforme apresentado no levantamento bibliográfico com as restrições e variáveis descritas no item anterior, os coeficientes de silhueta (SC) obtidos a partir da minimização da distância dentro de cada cluster estão na tabela 2.

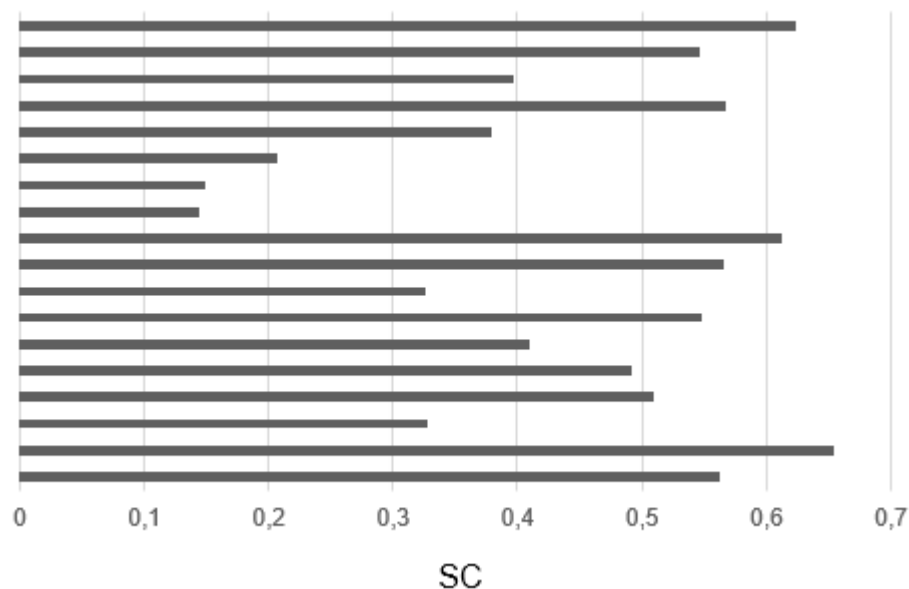
Tabela 2: coeficiente de silhueta das clusterizações testadas

Número de clusters	SC
3	0,45
4	0,33
5	0,40

Fonte: elaborado pela autora.

O coeficiente de silhueta pretende mensurar a proximidade dos pontos internos a um mesmo cluster assim como a distância deles aos demais clusters obtidos, garantindo que os segmentos sejam coesos e separados. Nesse sentido, um coeficiente maior implica uma clusterização mais coerente, obtida por meio da utilização de  $k=3$  no caso estudado. O gráfico 12 apresenta a distribuição das silhuetas do método selecionado, destacando a ausência de coeficientes negativos. Isso aponta para um resultado mais robusto, de forma que todos os pontos foram assignados ao cluster cujos pontos são mais próximos dele; nenhum objeto foi classificado erroneamente.

Gráfico 12: distribuição da silhueta para 3 clusters



Fonte: elaborado pela autora.

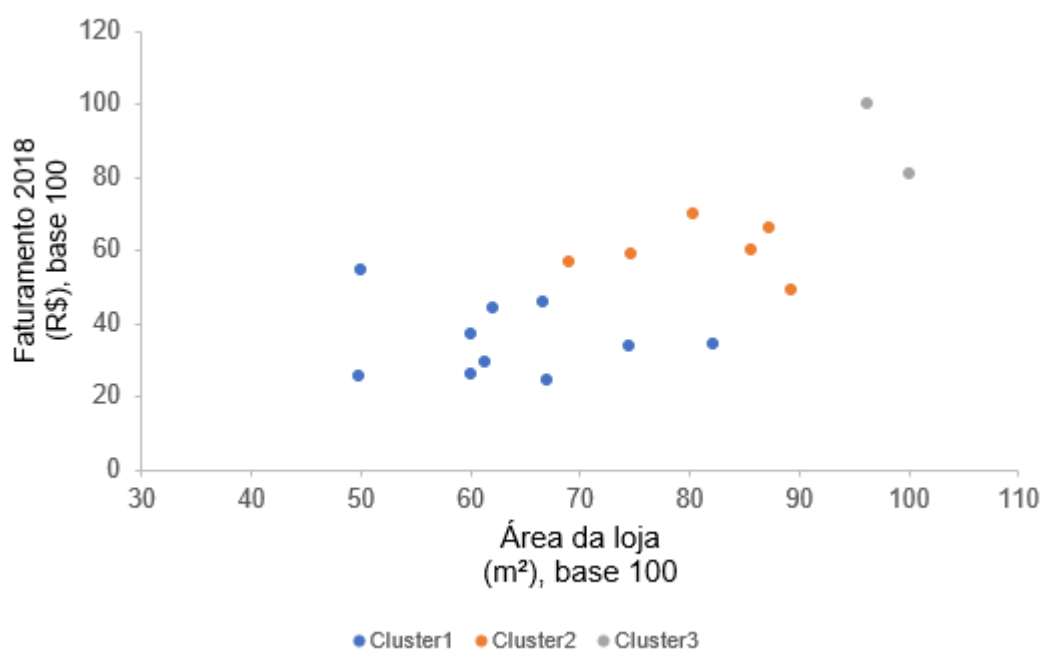
Os segmentos obtidos pelo método selecionado foram representados graficamente, de acordo com as variáveis utilizadas, em base 100, no gráfico 13. A tabela 3 apresenta os centroides dos clusters, também em base 100.

Tabela 3: centroides dos clusters

	Cluster 1	Cluster 2	Cluster 3
Área (m <sup>2</sup> )	62,76	82,36	99,14
Faturamento (R\$)	34,51	61,05	85,30

Fonte: elaborado pela autora

Gráfico 13: representação gráfica dos cluster obtidos



Fonte: elaborado pela autora

#### 4.5 Características dos segmentos obtidos

Os dois primeiros clusters foram posteriormente divididos em grupos de acordo com a classe social do bairro das lojas, obtida no site do Instituto Brasileiro de Geografia e Estatística (IBGE). A informação foi incluída na base de dados das lojas da Varejo SA assim como na planilha de clusterização, automatizando a obtenção da divisão final de clusters exibida no gráfico 14.

Essa informação é utilizada, sobretudo, para a redefinição do sortimento de produtos, cuja metodologia será explicada no item 4.6. Por esse motivo, apenas os dois primeiros clusters iniciais foram divididos por classe. Vale notar que, apesar das intenções de expansão e diferenciação das lojas da rede, a empresa tem uma variedade de produtos e preços que atende a um público amplo, mas relativamente bem definido. É válido assumir, portanto, que as futuras

unidades serão abertas em bairros de classes semelhantes às do mix atual de lojas e utilizar apenas as duas classificações: AB e BC.

- Cluster 1-AB: lojas pequenas, faturamento baixo, classe AB;
- Cluster 1-BC: lojas pequenas, faturamento baixo, classe BC;
- Cluster 2-AB: lojas grandes, faturamento médio, classe AB;
- Cluster 2-BC: lojas grandes, faturamento médio, classe BC;
- Cluster 3: lojas de área muito grande e faturamento elevado.

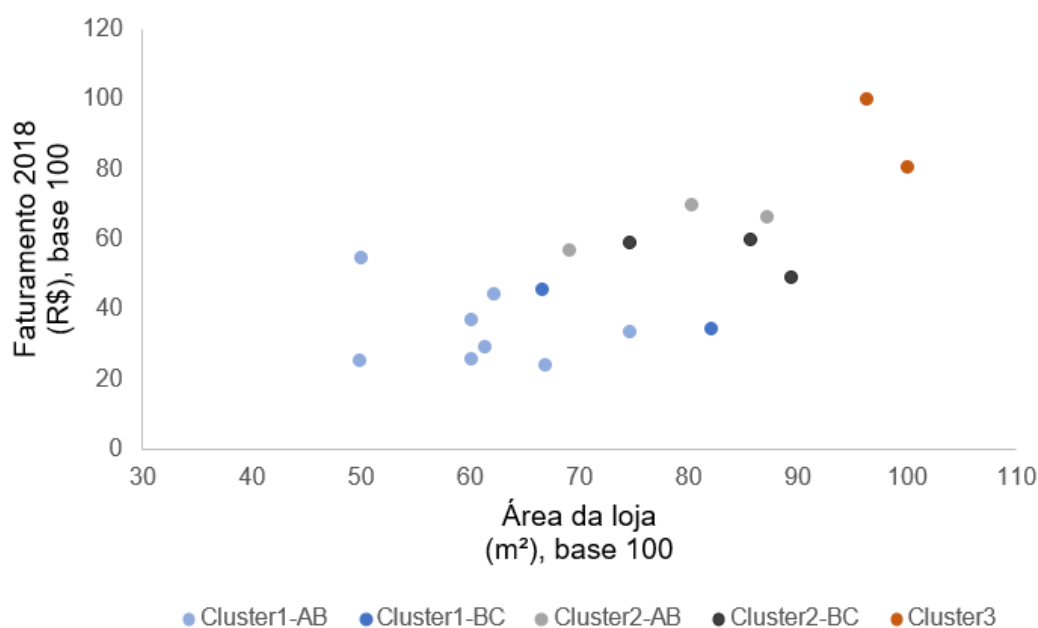
Os esforços de edição de sortimento devem ser especialmente focados nos dois segmentos obtidos a partir do cluster 1, de forma a reduzir a densidade de exposição, assim como melhorar sua performance. Outras estratégias de marketing direcionadas podem ser desenvolvidas para esse grupo, como treinamentos da força de vendas para melhorar a experiência dos consumidores e alavancar as vendas.

Os dois próximos segmentos são compostos por lojas de área maior e faturamento médio. Nesse caso, a redução do sortimento torna-se uma questão menos prioritária, já que sua performance e densidade estão em linha com as esperadas. Ainda assim, podem ser feitas algumas alterações menores e priorizações, além de outras iniciativas de marketing, com o intuito de elevar o faturamento do grupo.

O último cluster, que apresenta apenas dois representantes na atual conjuntura, é composto por lojas de grande porte – tanto em termos físicos quanto em receita. Devido à ampla área e boa performance, é ideal para experimentação da venda de novos produtos e apostas mais ousadas da empresa.

As variáveis analisadas para realização da segmentação sugerem que os segmentos obtidos são aceitáveis de acordo com os critérios de Blythe (2005). Os grupos são mensuráveis e têm características estáveis, coerentes e bem distintas, a partir das quais podem ser tomadas decisões bem informadas para cada um deles. Apesar de alguns segmentos apresentarem poucos pontos, podendo ser discutida sua substancialidade, é importante notar que, relativamente à pequena amostra atual de lojas, seu tamanho é justificado.

Gráfico 14: representação gráfica dos clusters finais



Fonte: elaborado pela autora.

#### 4.6 Redefinição do sortimento

Um dos principais objetivos da segmentação realizada, conforme explicado anteriormente, foi a edição do sortimento de produtos enviado a cada grupo de lojas. A consultoria contratada realizou um estudo aprofundado das categorias, linhas e, finalmente, dos produtos oferecidos pela Varejo SA em conjunto com o time de marketing da empresa, em duas grandes sessões de trabalho. Algumas definições envolviam grande parcela de subjetividade, sendo necessária a expertise de mercado do time de marketing para que as divisões fossem realizadas.

O portfólio completo de produtos regulares da empresa foi dividido em grupos de forma a facilitar o processo de definição do sortimento, os quais serão abaixo apresentados. O quadro 1 apresenta a definição final do sortimento de cada cluster, de acordo com a metodologia de divisão em “pacotes” criada. Além desses grupos, foram identificados com ajuda do time de marketing os produtos que não deveriam ser enviados para o varejo independentemente do segmento de lojas, por não serem condizentes com o posicionamento e imagem da marca para esse canal. Foram também retirados da lista de produtos alguns itens que haviam sido descontinuados mas continuavam ativos na base por terem estoque excedente ou por falha na gestão do portfólio, os quais apenas poluíam a base de produtos da empresa.



### Core

Conjunto de produtos indispensável, considerado o cerne do portfólio da Varejo SA. Foi definido de forma a englobar todos os itens de cada categoria classificados como A ou B na curva ABC, ou seja, aqueles responsáveis por 75% da receita das lojas. Além deles, foram incluídos outros itens das mesmas linhas de produtos, apesar de seu desempenho inferior, com os quais podem ser compostos conjuntos.

Para exemplificar, suponha que duas das categorias comercializadas pela Varejo SA são brincos e colares. Caso um modelo de brinco de ouro seja considerado *core* pela sua demanda elevada, o colar de modelo semelhante (mesma linha, apesar de outra categoria) deve também ser incluído no grupo, mesmo que tenha desempenho pior, de forma que a linha não seja desfalcada e o cliente possa compor o conjunto, elevando o valor da compra.

### Pacotes de Extensão de linha (AB ou BC)

Além do pacote *core*, foram criados dois pacotes de extensão de linhas, diferenciados pelo preço dos produtos. São incluídas linhas de produtos que não haviam sido priorizadas, definidas de acordo com a classe social do segmento.

A intenção desses pacotes é de oferecer um reforço do portfólio das lojas, para que o consumidor possa escolher a partir de uma variedade maior, mas de forma direcionada. Ou seja, ao invés de disponibilizar todos as linhas da empresa – possivelmente causando confusão – são oferecidas aquelas que têm mais chance de se moldar às necessidades dos clientes daquele grupo de lojas.

### Pacote de profundidade (AB ou BC)

O pacote de profundidade é definido por aqueles produtos complementares das linhas *core* e de extensão (a depender de qual delas foi selecionada, AB ou BC), que não foram incluídos nesses grupos devido à sua baixa performance. Diferente de outros produtos percebidos como essenciais na composição de conjuntos, eles não foram identificados dessa forma e acabaram sendo excluídos dos pacotes anteriores.

Quadro 1: pacotes incluídos no sortimento de cada segmento

	Core	Expansão AB	Expansão BC	Profundidade AB	Profundidade BC
Cluster1-AB	X	X			
Cluster1-BC	X		X		
Cluster2-AB	X	X		X	
Cluster2-BC	X		X		X
Cluster3	X	X	X	X	X

Fonte: elaborado pela autora

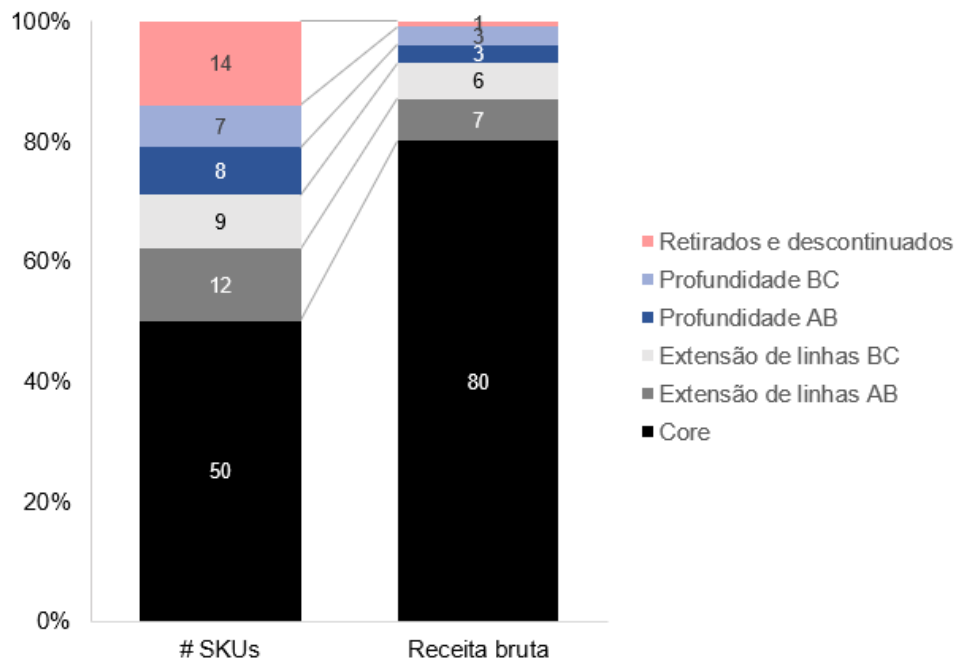
É importante ressaltar que esses pacotes dizem respeito apenas aos itens regulares do portfólio da empresa. As linhas virtuais e kits de datas comemorativas são enviados a todas as lojas, independentemente de seu segmento. Isso ocorre pois os esforços publicitários e midiáticos contribuem positivamente para a demanda deles, cujas vendas são comparáveis ou superiores às daquelas de itens A, historicamente.

#### 4.7 Resultados

A definição de grupos de produtos resultou na divisão do portfólio, em termos de número de SKUs e receita dos itens, apresentada no gráfico 15. Uma parcela significativa dos itens foi retirada por falta de aderência ao posicionamento da Varejo SA no canal estudado ou por descontinuação, mas que representava apenas 1% da receita das lojas. Os demais grupos definidos são também apresentados, sendo notável a relevância dos produtos *core* na receita do canal; apesar de totalizarem apenas metade dos itens disponíveis no portfólio, representaram historicamente 80% da receita do varejo.

Desconsiderando os itens retirados e descontinuados, foram cruzados os grupos de produtos enviados para cada segmento de lojas com a receita bruta que representaram historicamente nos respectivos segmentos. Essa análise demonstrou que o número de itens diferentes enviados para as lojas será reduzido em 9%, em média, chegando a mais de 20% de redução nas lojas dos segmentos 1-AB e 1-BC. Os itens removidos representavam menos de 3% da receita bruta dos segmentos dos quais foram retirados. Espera-se, no entanto, que essa receita perdida seja compensada e possivelmente superada pela compra de produtos similares aos removidos com maior aderência ao público do segmento.

Gráfico 15: relevância dos grupos de produtos definidos em número de itens e receita bruta do varejo



Fonte: elaborado pela autora a partir de dados da Varejo SA.

No que diz respeito à densidade de exposição, ou seja, o número de SKUs diferentes expostos por m<sup>2</sup> de área da loja, os resultados obtidos também foram bastante positivos. A redução do sortimento diminuiu a densidade média das lojas dos segmentos 1-AB e 1-BC em aproximadamente 25%, tornando-a comparável àquela identificada nas lojas de concorrentes, próxima a 20 SKUs/m<sup>2</sup>. Os demais segmentos têm suas densidades também reduzidas, mesmo o cluster 3 – que recebe todos os grupos de produtos – devido à remoção dos itens pouco atrativos e descontinuados.



## 5 PREVISÃO DE DEMANDA NA VAREJO SA

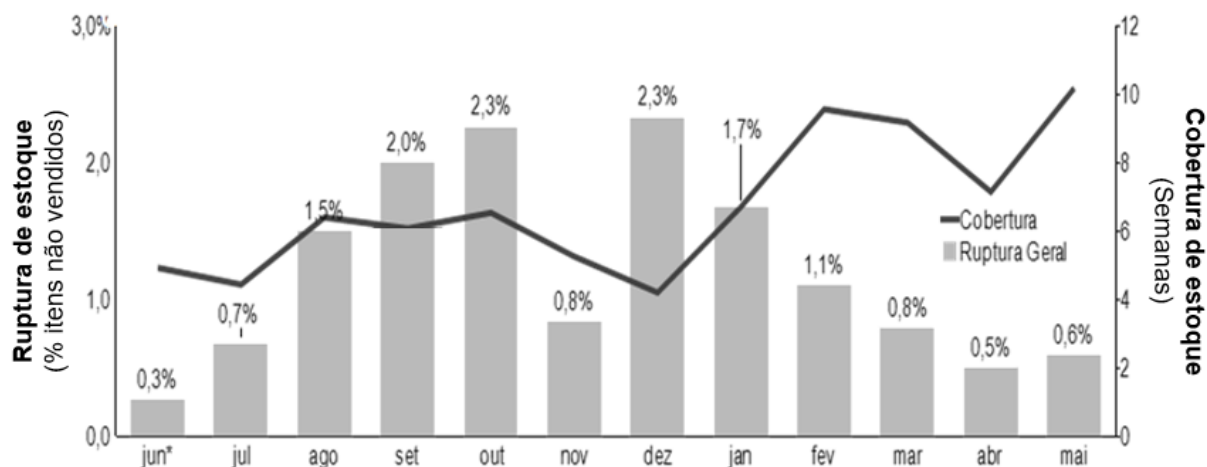
Nesse capítulo será brevemente apresentado o processo de previsão de demanda proposto em busca de atender melhor às necessidades da Varejo SA, demonstrando todo o procedimento de análise, construção do modelo e resultados.

O plano de produção do varejo da empresa é realizado usando como insumo a previsão de demanda que será revista nesse trabalho. Atualmente, a falta de acurácia desse processo leva a rupturas de estoque por falta de itens de alto giro no centro de distribuição e consequente tentativa de compensação pela oferta de itens cuja demanda foi superestimada, apresentando estoque excessivo. Um ponto a ser destacado é que a ruptura – ou seja, a falta de um produto que possui demanda – é altamente indesejável, pois prejudica a experiência do consumidor e frequentemente leva à não realização da venda. Dada a dimensão atual da operação do varejo, a falta de produtos de alto giro no estoque central do canal é muitas vezes compensada pela obtenção de itens excedentes nos estoques, consideravelmente maiores, de outros canais. No médio prazo, entretanto, essa medida não poderá ser utilizada devido ao crescimento esperado do varejo, cujas necessidades de estoque devem superar o excedente de outros canais.

No final de 2018, a empresa sofreu com um nível de ruptura elevado em suas lojas e, dado o processo incipiente de previsão de demanda existente, a área de planejamento comercial escolheu, como medida paliativa, aumentar os níveis de estoque de todos os produtos disponíveis nas lojas da marca. Assim, como mostra o gráfico 16, a falta de planejamento por trás da iniciativa de fato reduziu os níveis de ruptura, relacionados a itens de demanda elevada puxados dos estoques de outros canais, mas criou uma cobertura muito elevada para os SKUs em geral, devido ao acúmulo de itens de demanda baixa, que subiu de 5 para 10 semanas entre janeiro e junho de 2019.

Antes do projeto, a previsão de demanda na empresa era feita por meio de um software especializado de previsão da empresa *Analytic Always* ®. Devido à entrada de dados não tratados, à falta de parametrização adequada dos modelos quantitativos e ao nível de agregação inadequado para os propósitos da previsão, ela era muito imprecisa. Quando o projeto foi iniciado, a empresa já havia deixado de utilizar o software e solicitou que a nova ferramenta de previsão utilizasse apenas softwares já disponíveis, sem necessidade de investimento em um novo programa estatístico.

Gráfico 16: gráfico de ruptura e cobertura de estoque, Jun/18-Mai/19

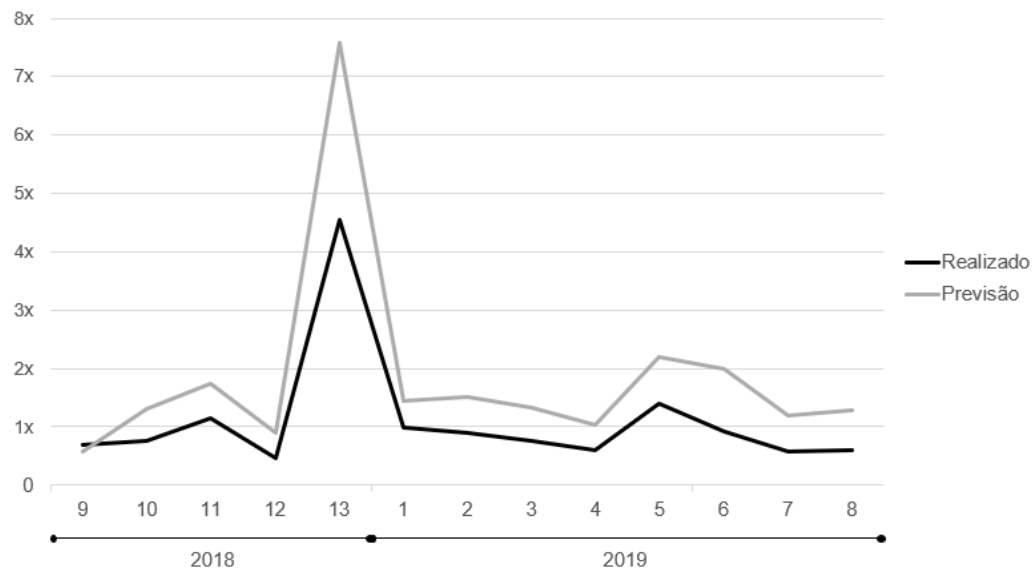


Fonte: elaborado pela autora a partir de dados da Varejo SA.

Foram extraídas as previsões históricas agregadas do software anteriormente utilizado e comparadas aos dados reais de vendas por campanha – divisão temporal que será explicada no item 5.1. O gráfico 17 apresenta a previsão de vendas do software e as vendas reais para o ano móvel de agosto de 2018 a julho de 2019. A tabela 4, que segue o gráfico, apresenta o erro percentual de cada período representado.

Pela análise do gráfico e dos erros percentuais por campanha, é possível identificar um forte viés da previsão, que em doze das treze campanhas analisadas calculou um valor de demanda superior ao real executado. Alguns períodos tiveram erros percentuais extremamente elevados, que impactam de forma significativa o planejamento da empresa, como aqueles das últimas três campanhas apresentadas; todos superiores a 100% da demanda real, em termos absolutos. Quando analisados os erros acumulados da previsão anual, os resultados também foram alarmantes: o erro percentual médio (MPE) foi de -70% e o erro percentual médio absoluto (MAPE) de 73%. Comprova-se, dessa forma, que a previsão apresenta forte viés assim como erro absoluto também muito alto.

Gráfico 17: gráfico de demanda realizada e prevista, Ago/18 a Jul/19



Fonte: elaborado pela autora a partir de dados da Varejo SA.

Tabela 4: erros percentuais da previsão do software, campanha 2018.09 a 2019.08

Campanha	Erro percentual
2018.09	19%
2018.10	-70%
2018.11	-53%
2018.12	-93%
2018.13	-67%
2019.01	-46%
2019.02	-69%
2019.03	-73%
2019.04	-70%
2019.05	-58%
2019.06	-116%
2019.07	-104%
2019.08	-115%

Fonte: elaborado pela autora a partir de dados da Varejo SA.

A seguir serão apresentados os objetivos da previsão de demanda, de acordo com o processo de planejamento adotado pela empresa. Dessa forma, será esclarecida a necessidade de se melhorar a acurácia da previsão, permitindo que a equipe tome decisões bem informadas para o planejamento de períodos futuros.

## 5.1 Propósito da previsão

Como mencionado no levantamento bibliográfico, Makridakis, Wheelwright e McGee (1998) destacam como primeiro passo para o desenvolvimento de um modelo de previsão a definição clara do objetivo ou propósito dela. No caso da Varejo SA, assim como na maior parte das empresas de manufatura ou varejo, a previsão da demanda se faz necessária em diferentes níveis de tomada de decisão – estratégico, tático e operacional.

A gestão do varejo é feita por campanhas, totalizando treze ciclos ao longo do ano, com diferentes durações entre si (geralmente, entre 3 e 5 semanas). Isso acontece porque, em cada campanha, são colocados certos produtos em destaque na mídia e lançados itens virtuais específicos ou kits de datas comemorativas (conforme apresentado na seção 1.2.3 Estrutura de Produto), de acordo com a oferta de itens planejada para aquele período em especial. Os únicos itens presentes ao longo de todo o ano no portfólio das lojas são os regulares, cuja demanda pode ser afetada pela oferta de produtos virtuais ou kits de data que desempenhem função semelhante.

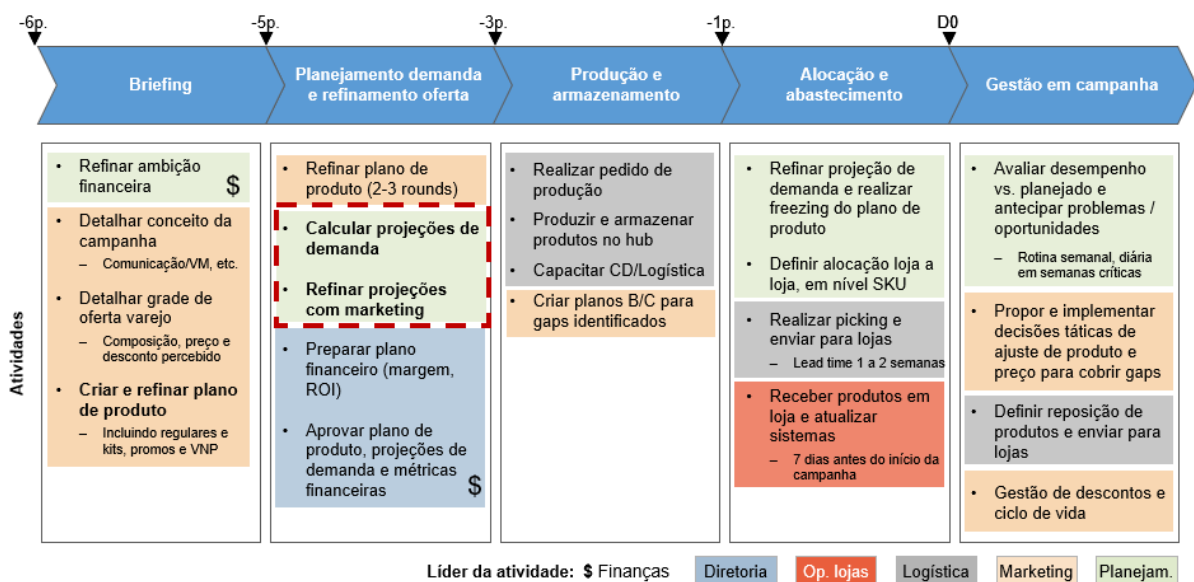
De acordo com as diretrizes de S&OP apresentadas na revisão bibliográfica, o plano estratégico da Varejo SA é revisto com regularidade, sendo realizadas reuniões de planejamento para cada campanha (treze ao ano). O planejamento de uma campanha da Varejo SA é composto por 5 etapas, cujas atividades e departamentos responsáveis (ou líderes) são detalhados na figura 13. O processo começa com aproximadamente 6 meses de antecedência, quando o departamento de marketing define os produtos que serão comercializados na campanha, assim como seus preços e estratégias de mídia, apresentados por meio de um documento chamado plano de produto. Após discussões com os times de planejamento e finanças sobre o plano de produto, ele é revisado e o time de planejamento calcula a demanda prevista para o período. Feita a previsão, os resultados são apresentados em fórum e discutidos levando em consideração os inputs do time de marketing, que tem conhecimento sobre tendências de mercado, apostas em produtos específicos etc. Ajustados os números, as projeções são aprovadas pelo departamento financeiro, de acordo com as metas estabelecidas, podendo ser realizado o pedido de produção 3 períodos antes do início da campanha. A produção dos itens leva de seis a oito semanas e o transporte até as lojas chega a tomar duas semanas. Com *lead time* máximo dos produtos sendo de 10 semanas, a produção deve ser iniciada com 3 campanhas de antecedência (12 semanas, em média), restando quase duas semanas para organizar o estoque e preparar as lojas para a nova campanha, além de lidar com eventuais contratemplos. Identifica-se, assim, a necessidade de previsão da demanda de uma campanha 3 períodos antes de seu início,



utilizando os dados disponíveis até a última campanha finalizada, para atualização do plano de produção e colocação de pedidos para a fábrica.

Uma vez concluída a produção, é realizada a alocação dos produtos para início de campanha, loja a loja. Em decorrência, os lotes são enviados às unidades correspondentes e os sistemas de estoque e ponto de venda atualizados com os produtos que estão planejados para aquela campanha. Por fim, a performance da campanha é gerenciada por meio de relatórios frequentes, que afetam as diretrizes de reabastecimento, promoções e implementação de táticas de ajuste de portfólio, quando necessário.

Figura 13: etapas e atividades de planejamento de campanha



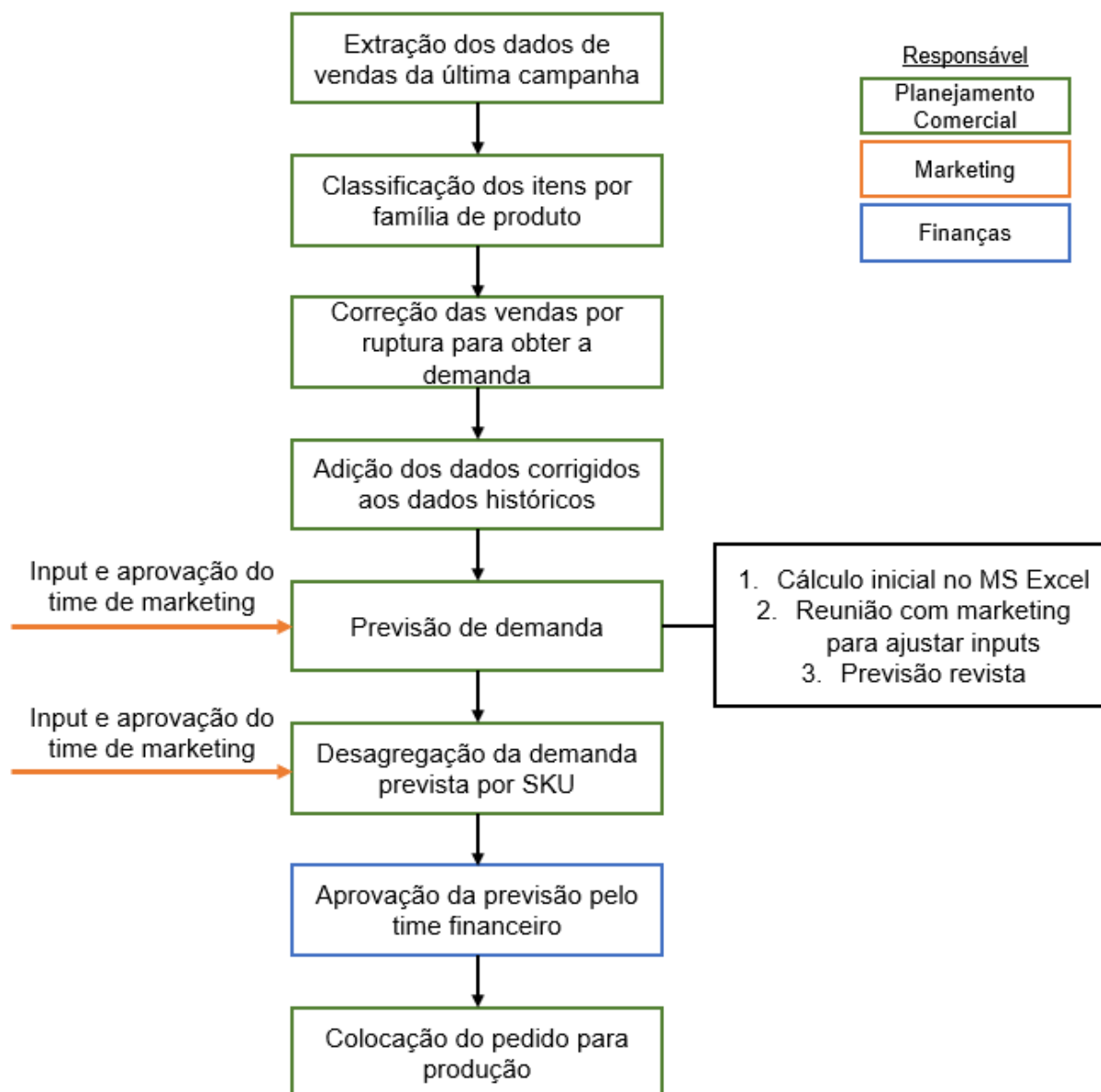
Fonte: elaboração da autora.

## 5.2 Procedimento de previsão proposto

Para garantir que a previsão da demanda incorpore os inputs necessários e os dados sejam tratados da maneira devida, foi desenhado um processo a ser seguido pelo time de planejamento. Ele deve ser realizado nos primeiros quatro dias de cada campanha, quando os resultados da anterior já estão disponíveis no sistema, com o objetivo de se prever a demanda de três campanhas a frente, para atualização do plano de produção. Alguns detalhes do processo serão mais bem explicados nos próximos itens, como a agregação por famílias e correção das vendas.

A figura 14 apresenta as atividades envolvidas no processo de previsão, assim como o time responsável por elas.

Figura 14: etapas do processo de previsão proposto



Fonte: elaboração da autora

### 5.3 Nível de agregação da previsão

Relembrando a estrutura de produtos da empresa apresentada no capítulo 1, a oferta de SKUs em uma campanha é composta por itens regulares (sempre disponíveis no portfólio), produtos virtuais disponíveis por um período específico e, no caso de datas comemorativas, os kits de datas vendidos apenas na campanha referente àquela comemoração. Além disso, após a implementação dos sortimentos personalizados definidos no capítulo anterior para cada segmento de lojas, a oferta de itens regulares também se tornou variável. Devido a essa alteração dos itens ofertados entre campanhas, foram considerados dois níveis de agrupamento

para cálculo da demanda: no nível SKU, procurando “espelhos” para os itens virtuais e kits ou no nível família, posteriormente segmentado pelos SKUs ofertados na campanha em questão. Devido ao dinamismo do portfólio existe uma dificuldade de se designar itens de desempenho semelhante entre campanhas, o que levaria a uma previsão no nível SKU ainda menos confiável do que seria com um portfólio contínuo. Além disso, a variedade de SKUs disponíveis a cada campanha é muito elevada, podendo chegar a quase mil itens diferentes, cujo tratamento individualizado seria extremamente complexo, demandaria muito tempo e dificultaria o trabalho do time de planejamento. Foi decidido, portanto, que a previsão de demanda seria feita de forma agrupada por família. Essa decisão contribui ainda mais para a precisão do modelo quantitativo desenvolvido, pois conforme a literatura relevante, um dos princípios de previsões é que quando feitas em níveis agregados apresentam acurácia maior do que por SKU.

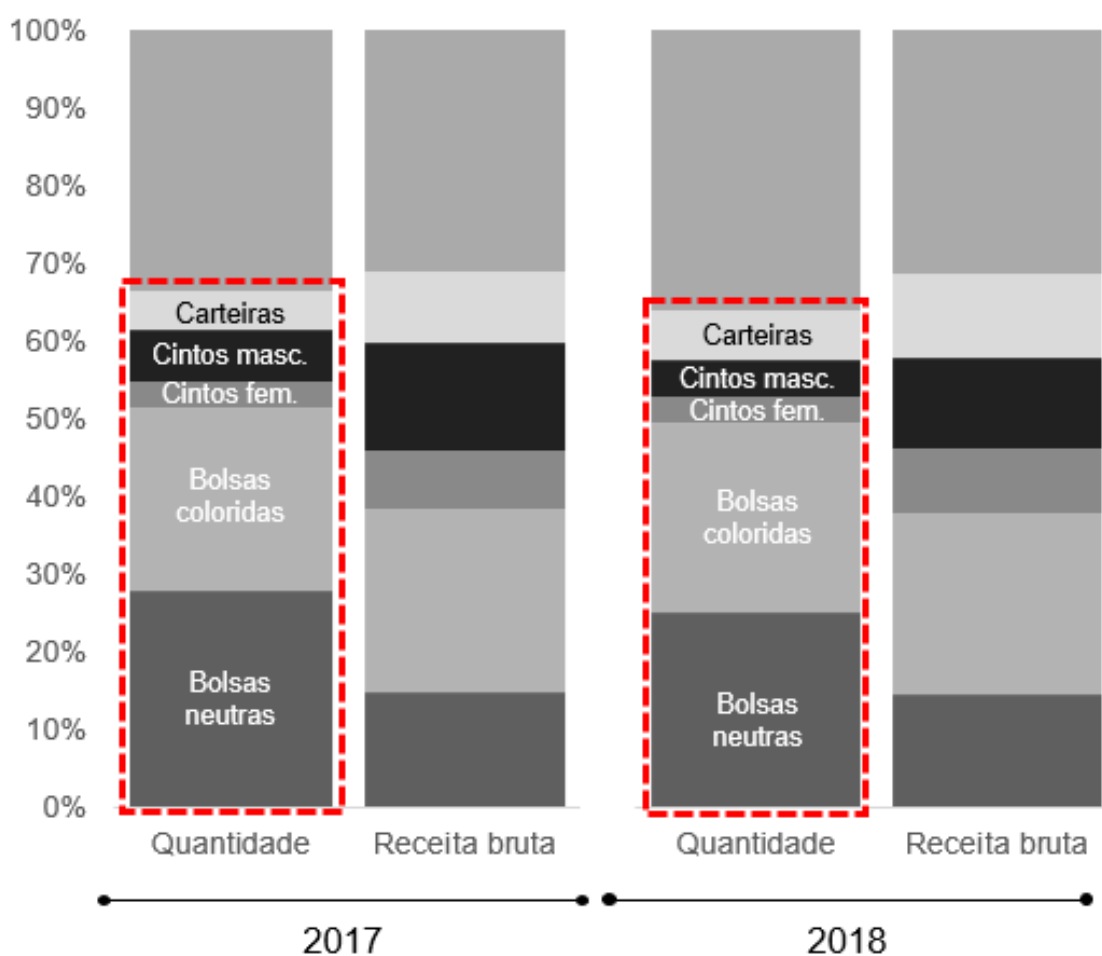
A definição das famílias de previsão foi realizada com o objetivo de agrupar itens com processos produtivos semelhantes e ao mesmo tempo semelhantes aos olhos do consumidor, cujas demandas sigam um padrão ao longo do tempo. Dessa maneira, é prevista a demanda total daquela família com base no desempenho histórico, de forma que a desagregação item a item possa ser realizada posteriormente de acordo com a oferta da campanha em questão, com base no desempenho histórico dos itens ou de produtos com características semelhantes a eles. Essa fase do processo, apesar de não ser tratada a fundo nesse trabalho, envolve um profundo conhecimento do portfólio da Varejo SA e do mercado, sendo necessário levar em conta as opiniões de especialistas do departamento de Marketing com relação às “apostas” da campanha, que podem variar conforme o plano de produto, e outras possíveis particularidades.

Como a indústria de atuação da Varejo SA não pode ser divulgada por questões de confidencialidade, suponhamos que seja do ramo de acessórios de moda femininos. Nesse caso, alguns exemplos de famílias de produtos estão elencados abaixo. Esses agrupamentos serão adotados no restante do trabalho para ilustrar a metodologia e facilitar a compreensão das análises decorrentes:

- Bolsas de tons neutros
- Bolsas coloridas
- Carteiras femininas
- Cintos femininos
- Cintos masculinos

No caso da Varejo SA, os itens foram repartidos em famílias que, conforme esperado, apresentaram comportamentos semelhantes ao longo dos anos completos para os quais existe histórico confiável de vendas (2017 e 2018). Para exemplificar a metodologia utilizada e resultados obtidos, serão apresentadas ao longo do trabalho as cinco famílias mais relevantes; nos dois anos totalizaram mais de 65% da quantidade de itens vendidos e 70% da receita bruta do canal, como demonstra o gráfico 18.

Gráfico 18: relevância das famílias no total de itens vendidos e receita bruta, 2017 e 2018.



Fonte: elaborado pela autora a partir de dados da Varejo SA.

#### 5.4 Coleta de dados

A obtenção de dados de qualidade, que serão usados para a construção do modelo de previsão, é muito importante para garantir sua acurácia. Não apenas os dados devem ser confiáveis como é preciso atentar para sua compatibilidade com o propósito da previsão, por meio da decisão do nível de agregação dos itens, períodos e realização de eventuais correções.

Os dados de vendas das lojas da Varejo SA são registrados em um software de gestão do ponto de venda provido pela Linx<sup>®</sup>, conectado diretamente ao sistema de *Enterprise Resource Planning* (ERP) da mesma desenvolvedora. O time de planejamento comercial tem total acesso às vendas de todas as lojas em um nível muito detalhado, por SKU dentro de cada ticket de venda.

Para atender aos objetivos da previsão apresentados anteriormente, foi decidido que os dados deveriam ser extraídos por semana, para realização de correções posteriormente mencionadas, e então agregados por campanha. Um nível de detalhe maior aumentaria a complexidade e quantidade de dados sem refletir em benefícios para a previsão, já que os objetivos delimitados não exigem maior detalhamento. Devido à extensão diversa das campanhas, foi realizado o ajuste da base conforme apresentado no item 3.8.2 Ajustes e transformações, substituindo-se a base mensal pela campanha. Dessa forma, os dados de demanda corrigidos ( $W_t$ ) foram obtidos a partir da seguinte equação:

$$W_t = Y_t \times \frac{\text{número de dias médio por campanha}}{\text{número de dias na campanha } t} = \frac{365,25/13}{\text{número de dias na campanha } t}$$

A extração foi feita por SKU, posteriormente classificados por família, conforme o procedimento apresentado na figura 14. Dessa forma, a previsão pode ser calculada diretamente no nível de agregação desejado para tomada de decisões. Foram utilizados dados de vendas de janeiro de 2017 até setembro de 2019, período no qual já existia uma quantidade relevante de lojas maduras (com 1 ano ou mais de operação), cujos dados são mais confiáveis para análise. As análises e previsões serão realizadas de acordo com a demanda média por loja, já que se fixado o número de lojas o volume de dados analisado seria pouco representativo.

Por fim, foram feitas algumas correções na base de dados para que as vendas refletissem de forma mais realista a demanda real das lojas da Varejo SA. Foi extraída uma base de estoques semanais das lojas, que apresenta o número de unidades disponíveis de cada SKU na noite anterior ao reabastecimento (realizado às quintas-feiras no caso da empresa estudada). Quando algum SKU terminava a semana sem estoque em loja, suas vendas semanais eram analisadas em busca de itens que sofreram ruptura – ou seja, casos em que aconteceu o não atendimento da demanda. Como aproximação, foi considerado que sofreram ruptura os itens sem estoque cujas vendas semanais foram inferiores à venda média semanal de itens da sua família – calculada considerando o histórico de vendas do ano tratado. Para que os dados refletissem a demanda real, reduzindo o impacto da ruptura de estoque, a venda do item na semana em que

sofreu ruptura foi aproximada pela venda média semanal de itens da família à qual pertence, calculada conforme a equação abaixo para cada uma das famílias estudadas:

$$Venda\ média\ semanal = \frac{\sum Q_{item, x}}{S \times I}$$

Onde:

- $Q_{item, x}$ : volume vendido (unidades) por item da família X no ano
- S: número de semanas no ano
- I: número de itens diferentes vendidos da família

Uma limitação encontrada para a realização dessa correção foi que a base de estoques semanais começou a ser disponibilizada apenas em junho de 2018, não sendo possível corrigir as vendas de itens anteriores a essa data.

## 5.5 Análise preliminar

Uma vez coletados os dados necessários, a próxima etapa do processo de previsão consiste na análise exploratória deles em busca da identificação de padrões e comportamentos que possam contribuir para a escolha dos métodos de previsão mais eficientes. Considerando que a previsão deveria ser feita em agrupamentos de itens (nível família) para cada uma das campanhas, a exploração dos dados foi realizada utilizando o mesmo padrão.

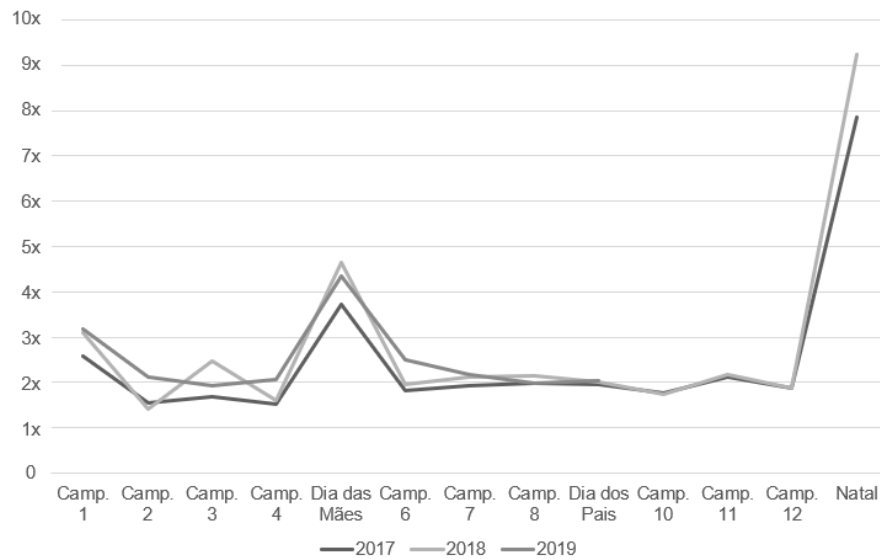
Além das análises abaixo apresentadas, foram utilizados gráficos de autocorrelação para comprovar a existência (ou ausência) de sazonalidade nas séries estudadas, os quais são apresentados no anexo A.

### Bolsas de tons neutros

As bolsas de tons neutros são uma das famílias com maior venda da Varejo SA, devido à versatilidade do produto assim como a sua “presenteabilidade”. Isso significa que esse tipo de produto tem volume de vendas normalmente alto, mas a facilidade de ser presenteado alavanca ainda mais suas vendas em algumas datas específicas.

Para confirmar esse efeito de sazonalidade da família em questão, foi desenhado o gráfico de sazonalidade da demanda a partir dos dados de vendas corrigidas por campanha de 2017, 2018 e 2019 (parciais), apresentado no gráfico 19.

Gráfico 19: gráfico de sazonalidade da demanda de bolsas neutras, Jan/2017-Ago/2019

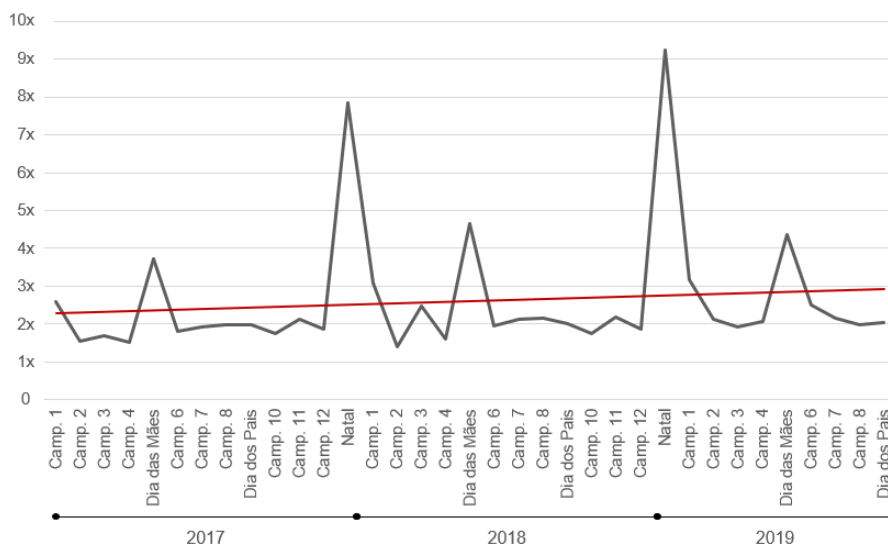


Fonte: elaborado pela autora a partir de dados da Varejo SA.

Fica evidente, a partir da comparação entre anos, o aumento da demanda da família em algumas datas comemorativas, assim como na primeira campanha do ano, que pode ser facilmente explicado. Pela natureza do produto em questão, bolsas femininas de tons neutros, é esperado que as vendas não sejam alavancadas no Dia dos Pais, mas sim em datas em que mulheres são presenteadas, tais quais Dia das Mães e Natal. A campanha 1, por sua vez, é aquela em que ocorrem as mais significativas promoções programadas nas lojas da Varejo SA, seguindo o calendário geral do varejo que costumeiramente apresenta descontos após o Natal.

A análise gráfica da série histórica de forma contínua não permite identificar uma forte tendência na demanda da família estudada. O gráfico 20 apresenta a demanda por campanha para o histórico disponível, assim como a tendência linear da série, em vermelho. A leve inclinação da tendência sugere um fator muito baixo, sendo possível que um método de previsão que não leve em consideração tal característica funcione para esse grupo de produtos.

Gráfico 20: demanda por campanha, Jan/2017-Ago/2019

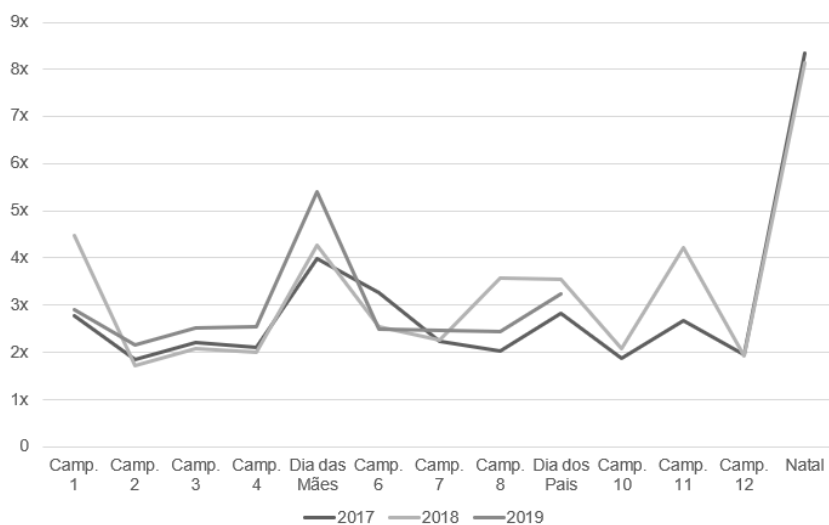


Fonte: elaborado pela autora a partir de dados da Varejo SA.

### Bolsas coloridas

Da mesma forma que as bolsas neutras, as coloridas são produtos destinados majoritariamente ao público feminino, cujas vendas são fortemente alavancadas por datas comemorativas e promoções, conforme apresentado no gráfico 21, de sazonalidade. É evidente o aumento da demanda consistente nos três anos nas mesmas três campanhas, a primeira por conta dos descontos e o dia das Mães e Natal devido à compra de itens como presentes.

Gráfico 21: gráfico de sazonalidade da demanda de bolsas coloridas, Jan/2017-Ago/2019

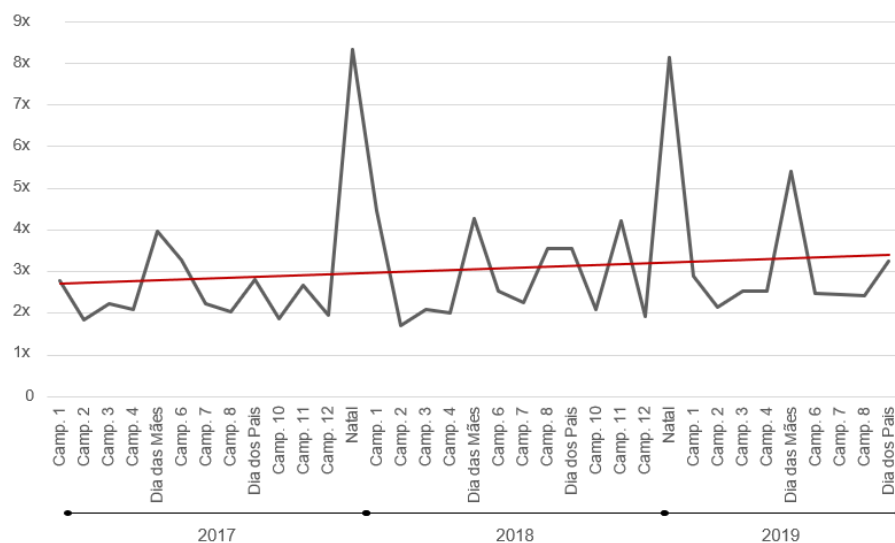


Fonte: elaborado pela autora a partir de dados da Varejo SA.



A comparação da demanda média semanal por campanha entre as famílias de bolsas neutras e coloridas revela comportamentos semelhantes devido à similaridade dos produtos. Conforme esperado pela similaridade com a família de bolsas neutras, novamente não fica evidente a existência de uma tendência de crescimento ou diminuição da demanda com o tempo. A análise do gráfico 22, de demanda por tempo, com a linha de tendência correspondente em vermelho, não permite concluir que exista tendência positiva ou negativa na demanda de bolsas coloridas.

Gráfico 22: demanda por campanha, Jan/2017-Ago/2019

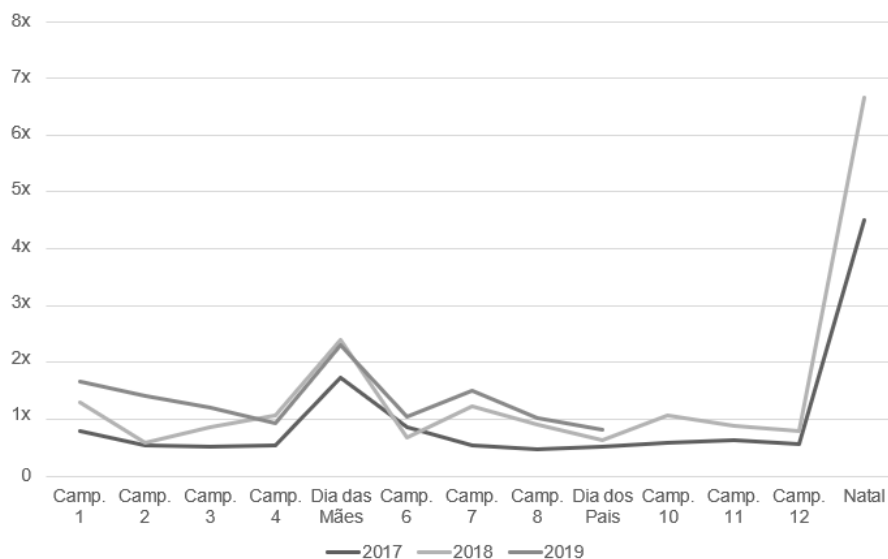


Fonte: elaborado pela autora a partir de dados da Varejo SA

### Cintos femininos

Os cintos femininos, de forma semelhante às bolsas, apresentam picos de vendas nas datas de comemorações que envolvem mulheres. Por outro lado, não fica claro em todos os anos um aumento da demanda pelos itens na primeira campanha, em que ocorrem as promoções. Ao consultar o time de marketing da Varejo SA quanto à essa particularidade, foi destacado que esse tipo de item costuma ser menos alavancado pelas promoções por relacionar-se a uma compra mais premeditada e de necessidade, não tão impulsiva quanto a de bolsas e outros tipos de acessórios. O gráfico 23, de sazonalidade, evidencia esses pontos.

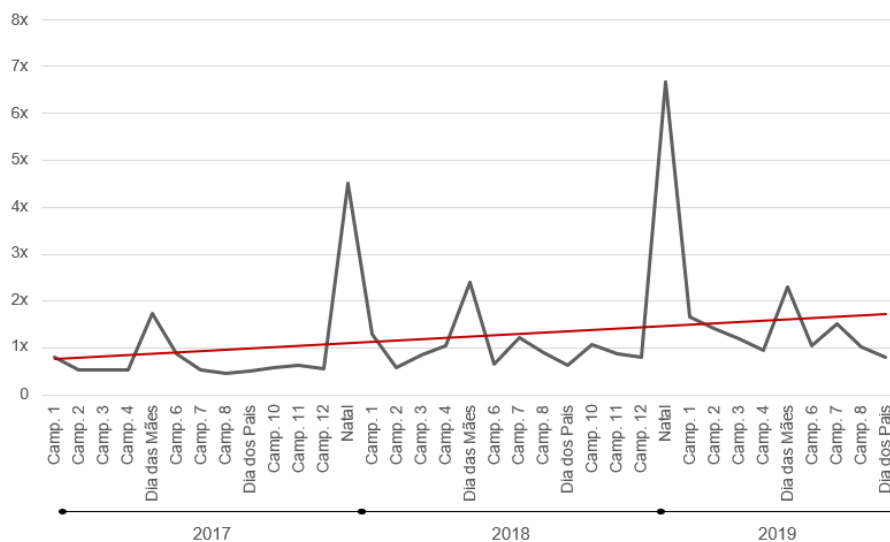
Gráfico 23: gráfico de sazonalidade da demanda de cintos femininos, Jan/2017-Ago/2019



Fonte: elaborado pela autora a partir de dados da Varejo SA.

Novamente, não é identificada forte tendência positiva ou negativa pela análise gráfica da série histórica, disponível no gráfico 24.

Gráfico 24: demanda por campanha, Jan/2017-Ago/2019



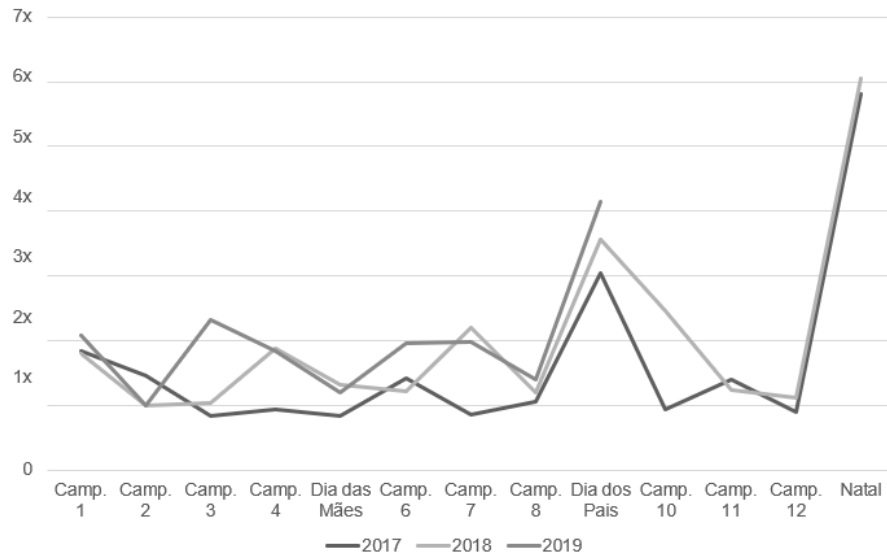
Fonte: elaborado pela autora a partir de dados da Varejo SA.

### Cintos masculinos

Os cintos masculinos, diferentemente das bolsas e cintos femininos, apresentam os picos de demanda no dia dos Pais e Natal, como esperado pela natureza do produto. Assim como os

outros cintos, por não serem um item de compra impulsiva, suas vendas não são alavancadas fortemente na temporada de descontos do início do ano, como evidente no gráfico 25.

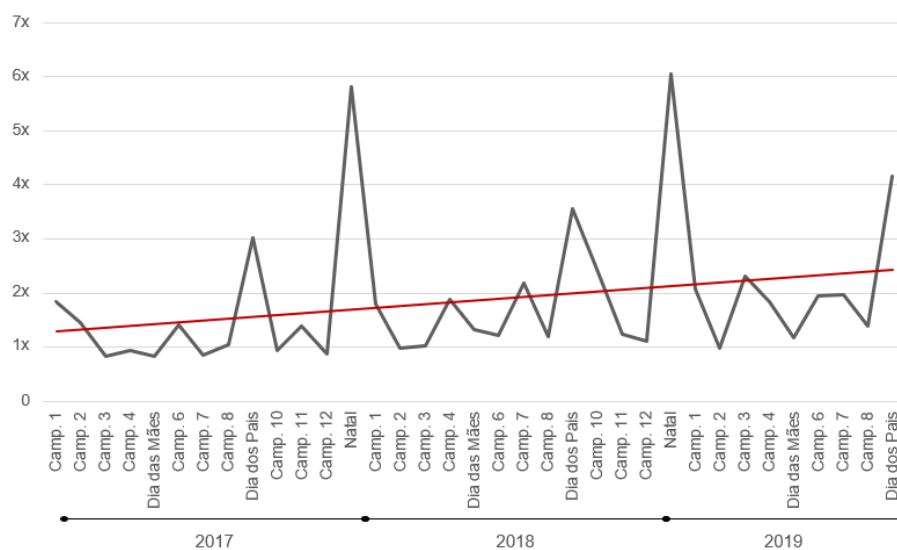
Gráfico 25: gráfico de sazonalidade da demanda de cintos femininos, Jan/2017-Ago/2019



Fonte: elaborado pela autora a partir de dados da Varejo SA.

A família em questão apresenta uma tendência linear levemente positiva, traçada em vermelho no gráfico 26, mas ainda pouco representativa.

Gráfico 26: demanda por campanha, Jan/2017-Ago/2019

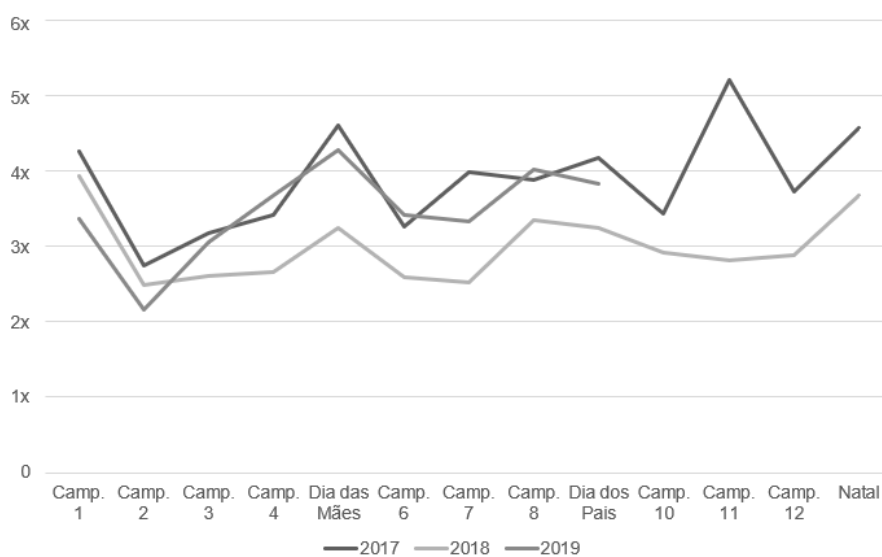


Fonte: elaborado pela autora a partir de dados da Varejo SA.

### Carteiras femininas

A sazonalidade das carteiras, diferentemente das demais famílias analisadas, não é tão relevante. Pelo gráfico 27, de sazonalidade, é perceptível um aumento mais discreto da demanda nas datas comemorativas relevantes para esse tipo de produto, assim como na primeira campanha de cada ano. Os desvios em relação à demanda média entre campanhas, entretanto, são menores do que aqueles apresentados para as demais categorias. Quando consultado em busca de explicações para essas particularidades, o time de marketing destacou que esses itens são de reposição mais frequente e menos presenteáveis do que as outras famílias analisadas, não tendo suas vendas alavancadas de maneira tão relevante pelas promoções ou datas comemorativas.

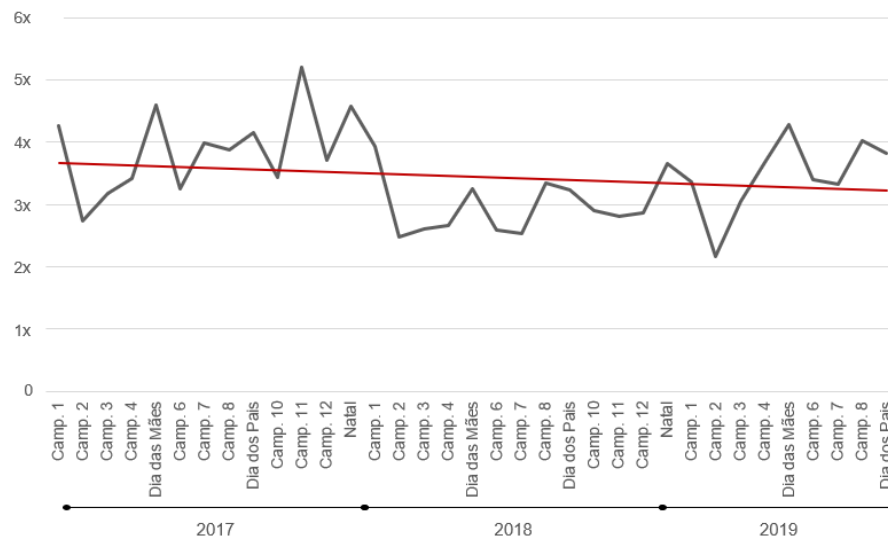
Gráfico 27: gráfico de sazonalidade da demanda de carteiras femininas, Jan/2017-Ago/2019



Fonte: elaborado pela autora a partir de dados da Varejo SA.

Assim como os demais produtos, a demanda das carteiras femininas não apresenta forte tendência. A linha vermelha do gráfico 28 apresenta a tendência linear da série, levemente negativa – diferentemente das outras famílias, com tendências também pouco representativas, mas com inclinação positiva.

Gráfico 28: demanda por campanha, Jan/2017-Ago/2019



Fonte: elaborado pela autora a partir de dados da Varejo SA

## 5.6 Escolha dos métodos utilizados

Uma vez realizada a análise exploratória e inspeção gráfica dos dados, foram comparados os métodos de previsão estudados no levantamento bibliográfico de forma a escolher aqueles que seriam utilizados no projeto. Com base nas aplicações de cada metodologia e as características das séries históricas, foi possível delimitar as opções para criação do modelo de previsão.

Foram analisados os quatro fatores de decisão estudados por Makridakis, Wheelwright e McGee (1998), apresentados no item 3.3.10 do presente trabalho, sendo eles: a natureza dos dados, características da série, tipo de dado (período) e frequência da previsão.

No caso da Varejo SA, o curto histórico de demanda disponível, devido à imaturidade do negócio do varejo, dificulta a realização de correlações robustas e confiáveis. Portanto, apesar de se saber que certos fatores afetam a demanda dos produtos, é difícil comprovar e calcular a relação exata com os dados existentes. Dessa forma, entende-se que é preferível a utilização de modelos de previsão por projeção das séries temporais.

Em relação às características da série, diversas informações sobre as famílias estudadas foram obtidas na inspeção gráfica realizada no item anterior desse capítulo. Foi percebida uma sazonalidade significativa nas famílias – exceto pelas carteiras femininas – não ficando clara nesse momento sua natureza aditiva ou multiplicativa. Por outro lado, não foi identificado forte padrão de tendência em nenhuma das famílias apresentadas, sendo interessante comparar-se os resultados de métodos que incluam esse componente com outros que não o façam.

A complexidade do método utilizado deve, ainda, respeitar o tipo de dado e frequência da previsão, assim como as restrições tecnológicas e organizacionais da empresa. Para os objetivos da Varejo SA, os dados são tratados por campanha e o processo deve ser repetido no início de cada uma delas, totalizando treze ciclos de previsão com um horizonte de três campanhas à frente. Deve-se levar em consideração que a empresa não pretende investir imediatamente em novas soluções automatizadas de previsão, como *Forecast Pro*® ou a solução de *Merchandise Planning* da *Analytic Always*® anteriormente utilizada. Além disso, o time de planejamento tem a previsão como uma de suas atividades prioritárias, mas acompanhada de diversas outras responsabilidades. Dessa forma, é esperado que a metodologia quantitativa desenvolvida seja robusta e acurada, utilizando softwares livres e com os quais o time está familiarizado, de forma que o processo não tome muitos dias de trabalho dos analistas responsáveis pela sua operação.

Considerando todas essas restrições de dados, histórico, características da série, frequência da previsão e comparando-as com os métodos estudados e suas aplicações indicadas, foram definidas as técnicas a serem testadas. Concluiu-se que serão metodologias estatisticamente robustas que levem em consideração a sazonalidade da demanda e possam ser facilmente compreendidas e utilizadas pelo time, uma vez construído o modelo. As técnicas selecionadas para testar as previsões de cada família foram:

- Suavização exponencial com sazonalidade aditiva
- Suavização exponencial com sazonalidade multiplicativa
- Suavização de Holt Winters com sazonalidade aditiva
- Suavização de Holt Winters com sazonalidade multiplicativa

## **5.7 Estrutura do modelo**

Conforme citado anteriormente, a Varejo SA não planejava investir em um novo software e buscava um modelo de previsão de demanda com resultados robustos, mas de simples compreensão e utilização pelos responsáveis. Para tanto, foi utilizado o software Microsoft Excel, que é a ferramenta de análise de dados mais usada na empresa e cujas principais funcionalidades eram dominadas pelo time de planejamento. Foram propostos outros softwares livres como a programação em R, mas o time de planejamento preferiu seguir com aquele que se familiarizava, pelo menos inicialmente.

O arquivo foi construído com uma aba de input de dados, na qual figuram os dados de demanda históricos e que deve ser atualizada com as novas demandas reais ao final de cada campanha. Para facilitar o processo de atualização, a planilha tem o mesmo formato que aquela extraída do software de ERP, bastando que o analista cole os dados corrigidos ao final da tabela. Uma ferramenta separada de correção dos dados por ruptura, conforme o procedimento explicado no item 5.4, também foi fornecida ao time.

Em outra aba, os dados de demanda são resumidos por família de produto, por meio de tabelas dinâmicas que devem também ser atualizadas a cada inserção de novos dados de demanda. Nessa mesma planilha, os valores de demanda por campanha são transformados pela extensão média, para que todas as campanhas sejam comparáveis. São esses os dados agregados e corrigidos que serão utilizados para calcular as previsões.

Por fim, são construídas três abas que realizam a previsão de demanda das famílias por cada um dos métodos escolhidos para teste. No topo da planilha existe um campo para seleção da família a ser prevista e há células para definição dos parâmetros relevantes em cada modelo, assim como uma tabela resumo dos erros analisados em cada previsão (MPE e MAPE). Abaixo são calculadas as previsões de acordo com os horizontes determinados (até 4 campanhas) e os erros individuais de cada período e família. Ao lado direito se encontra um gráfico que compara as demandas reais e previsões realizadas com 4 campanhas de antecedência, utilizadas para planejamento da produção.

Os dados de demanda disponíveis incluíam desde a primeira campanha de 2017 até a nona campanha de 2019, totalizando 35 períodos. Para testar a acurácia do modelo de acordo com previsões reais foi estabelecido um grupo de teste equivalente a 9 períodos, incorporando todo o ano de 2019. Dessa forma, o período de calibragem inclui os dois primeiros anos da série, equivalentes a 26 campanhas. As previsões para as últimas campanhas de 2019 que ainda não haviam ocorrido (10 a 13) também foram realizadas. A divisão da série histórica é apresentada na figura 15, na qual as campanhas são denominadas da seguinte forma: *Ano.Campanha*.

Figura 15: divisão dos dados históricos para construção do modelo



Fonte: elaborado pela autora

Makridakis, Wheelwright e McGee (1998) afirmam que um procedimento eficiente para inicializar os parâmetros de suavizações exponenciais ( $\alpha$ ,  $\beta$ ,  $\gamma$ ) é de usar valores elevados, que serão rapidamente ajustados e causarão menor impacto nas previsões. Seguindo esse procedimento, as constantes foram iniciadas e otimizadas, com auxílio da ferramenta Solver do Microsoft Excel, com o objetivo de minimizar o MAPE da previsão realizada para o grupo de inicialização. No que diz respeito aos valores iniciais dos índices utilizados em cada método, a literatura apresenta diferentes possibilidades de cálculo. Foram utilizadas as equações sugeridas por Makridakis, Wheelwright e McGee (1998):

- Nível:  $L_s = \frac{1}{s} (Y_1 + Y_2 + \dots + Y_s) = \frac{1}{13} (Y_1 + Y_2 + \dots + Y_{13}) =$
- Tendência:  $bs = \frac{1}{s} \left( \frac{Y_{s+1} - Y_1}{s} + \dots + \frac{Y_{s+s} - Y_s}{s} \right) = \frac{1}{13} \left( \frac{Y_{14} - Y_1}{13} + \frac{Y_{15} - Y_2}{13} + \dots + \frac{Y_{26} - Y_{13}}{13} \right)$
- Sazonalidade multiplicativa:  $S_1 = \frac{Y_1}{L_{13}} \quad S_2 = \frac{Y_2}{L_{13}} \quad \dots \quad S_{13} = \frac{Y_{13}}{L_{13}}$
- Sazonalidade aditiva:  $S_1 = Y_1 - L_{13} \quad S_2 = Y_2 - L_{13} \quad \dots \quad S_{13} = Y_{13} - L_{13}$

As previsões e índices para cada método foram calculados conforme as equações apresentadas no levantamento bibliográfico do capítulo 2.

## 5.8 Avaliação dos resultados

Os quatro métodos selecionados foram testados para cada uma das famílias, já que os padrões de consumo variam entre os tipos de produto de cada uma delas. Dessa forma, pode ser que um método seja mais adequado para bolsas coloridas, que apresentam fortes indícios de sazonalidade, enquanto outro se ajuste melhor à família de carteiras femininas, por exemplo.

Os resultados das previsões para o grupo de teste (*holdout set*) foram avaliados pelos valores de erro percentual médio absoluto (MAPE) calculado para cada família e método. Quando a diferença do MAPE entre métodos foi considerada irrelevante (inferior a 1%), optou-se por selecionar aquele que apresentava o menor erro percentual médio (MPE), ou seja, que era menos enviesado. As tabelas 5 a 9 apresentam os valores dos erros calculados para cada método e família. Foram incluídos também os erros das previsões antigas, realizadas pelo software da *Analytic Always*®, para termos de comparação.

Os gráficos apresentados para o método vencedor dizem respeito às previsões realizadas com quatro campanhas de antecedência para os grupos de inicialização, *holdout* e previsão, utilizando-se até o último dado disponível, da campanha 2019.09.



Para facilitar a leitura, os métodos foram denominados A, B, C e D, sendo eles respectivamente:

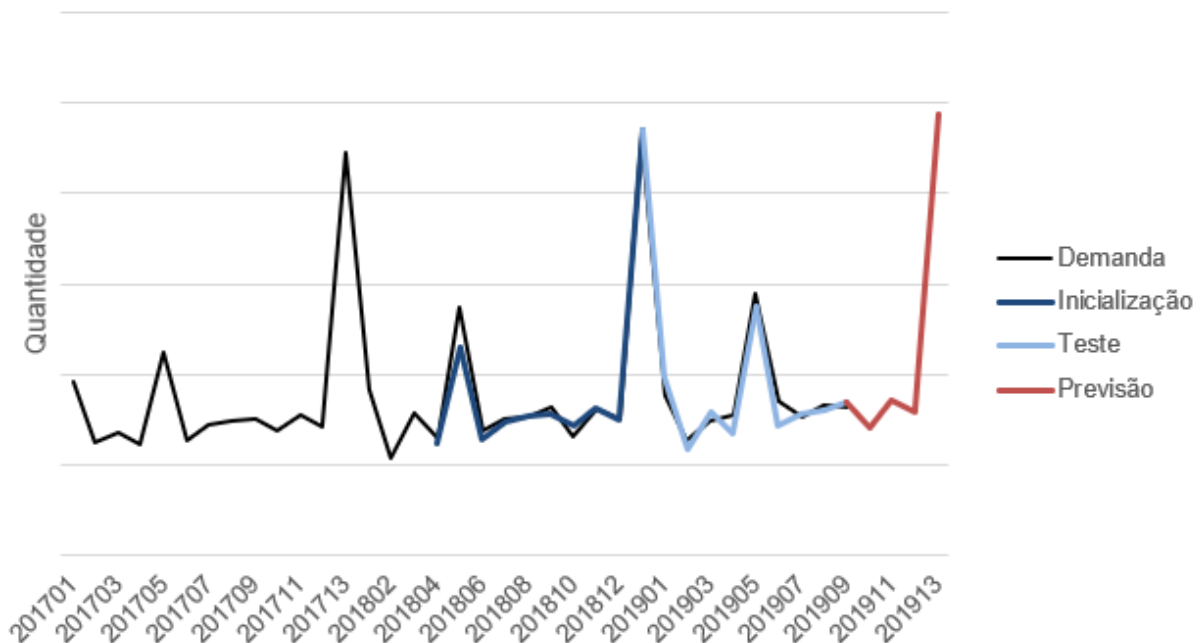
- A. Suavização exponencial com sazonalidade multiplicativa;
- B. Suavização exponencial com sazonalidade aditiva;
- C. Holt Winters com sazonalidade multiplicativa;
- D. Holt Winters com sazonalidade aditiva.

#### Bolsas de cores neutras

No caso das bolsas neutras, o método que apresentou maior acurácia para o *holdout set* foi a suavização exponencial de Holt Winter's com sazonalidade multiplicativa, cujas previsões são apresentadas no gráfico 29. Os métodos de suavização exponencial com sazonalidade também apresentaram resultados satisfatórios. Os coeficientes obtidos pela otimização do modelo foram  $\alpha = 0,02$ ,  $\beta = 0,4$  e  $\gamma = 0,7$ .

Em comparação à previsão antiga, o MAPE foi reduzido em 87%. O maior erro pontual, apresentado na campanha 2019.06 foi de 15,5%.

Gráfico 29: Previsão da demanda de bolsas de cores neutras



Fonte: elaborado pela autora.

Tabela 5: erros das previsões para bolsas de cores neutras

	MAPE	MPE
<b>A</b>	8,7%	4,6%
<b>B</b>	8,4%	8,1%
<b>C</b>	7,8%	2,4%
<b>D</b>	42,6%	45,2%
<b>Antiga</b>	63,6%	

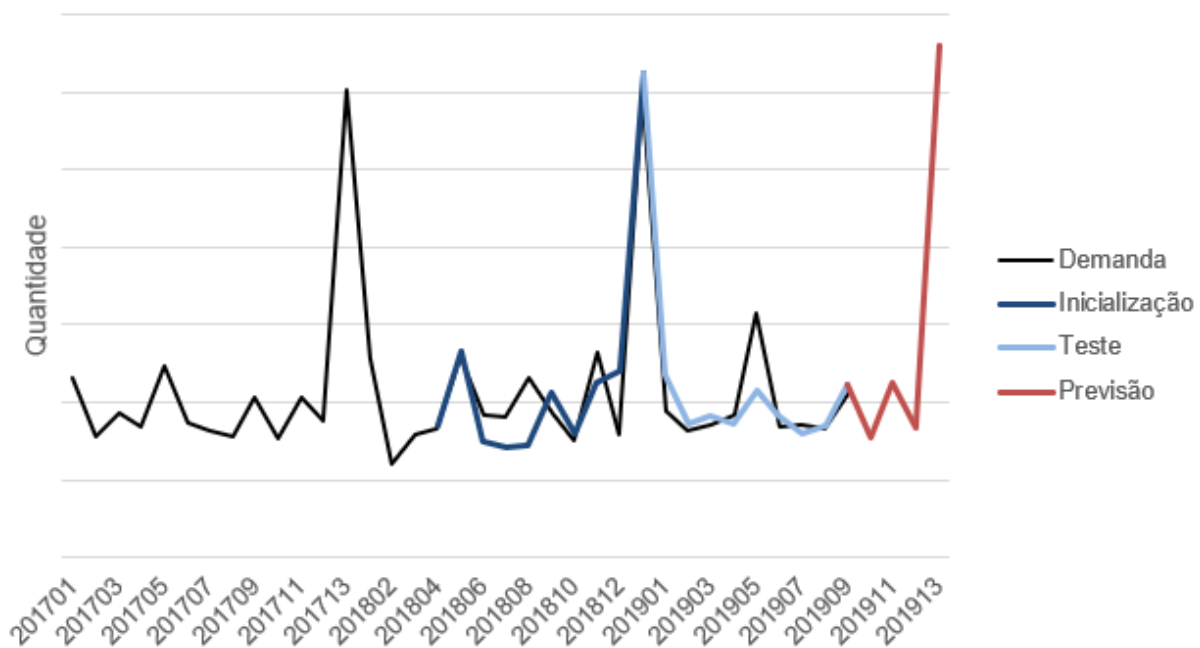
Fonte: elaborado pela autora.

### Bolsas coloridas

No caso das bolsas coloridas, que apresentaram um comportamento muito semelhante ao da primeira família na fase de análise exploratória dos dados, o método que apresentou menor erro foi a suavização exponencial com sazonalidade multiplicativa. Os coeficientes obtidos, a partir de otimização na fase de inicialização, foram  $\alpha = 0,71$ ,  $\gamma = 0,7$ .

O MAPE, em comparação ao da previsão antiga, foi reduzido em 85%. O maior erro pontual, apresentado na campanha 2019.05 foi de 31,5%.

Gráfico 30: Previsão da demanda de bolsas coloridas



Fonte: elaborado pela autora.

Tabela 6: erros das previsões para bolsas coloridas

	MAPE	MPE
<b>A</b>	11,0%	-0,8%
<b>B</b>	13,7%	13,9%
<b>C</b>	14,3%	-4,5%
<b>D</b>	36,2%	27,3%
<b>Antiga</b>	75,0%	

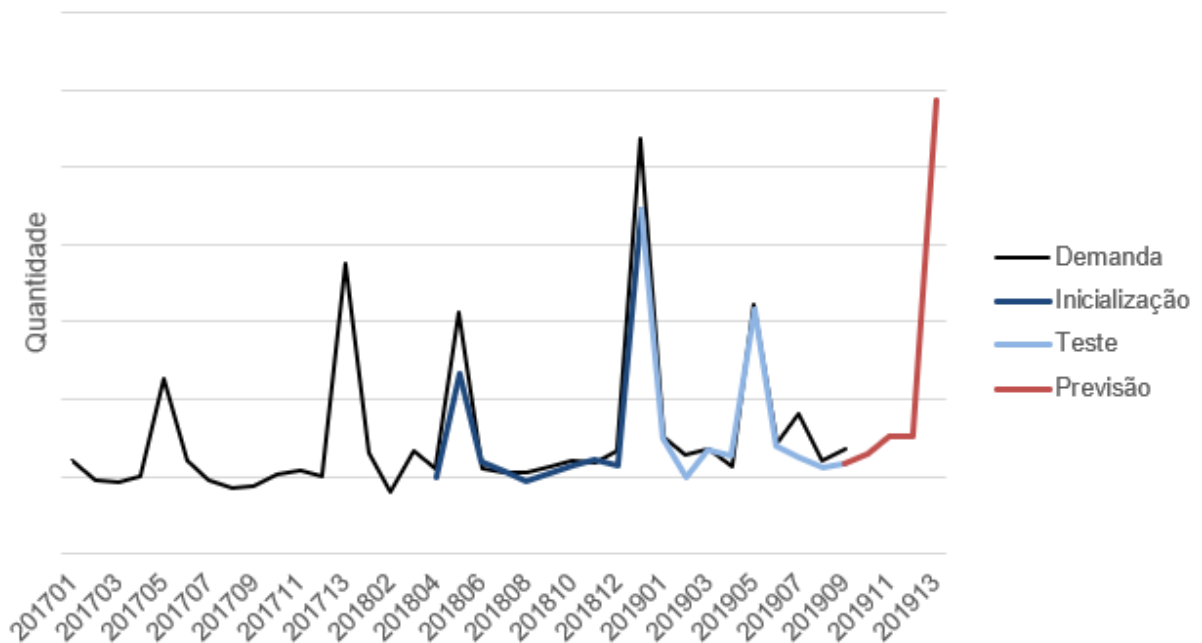
Fonte: elaborado pela autora.

### Cintos femininos

No caso dos cintos femininos, o método que apresentou menor MAPE foi a suavização exponencial com sazonalidade aditiva. No entanto, devido à diferença irrelevante desse erro em relação ao da suavização com sazonalidade multiplicativa (0,1%), essa foi selecionada por apresentar um MPE mais baixo. Os coeficientes utilizados foram  $\alpha = 0,29$  e  $\gamma = 0,7$ .

A redução do erro, em comparação ao modelo utilizado anteriormente pela empresa, foi de 84%. O maior erro pontual, apresentado na campanha 2019.07 foi de 31,7%.

Gráfico 31: Previsão da demanda de cintos femininos



Fonte: elaborado pela autora.

Tabela 7: erros das previsões para cintos femininos

	MAPE	MPE
<b>A</b>	10,5%	7,7%
<b>B</b>	10,4%	10,8%
<b>C</b>	13,0%	-4,1%
<b>D</b>	47,0%	34,6%
<b>Antiga</b>	64,4%	

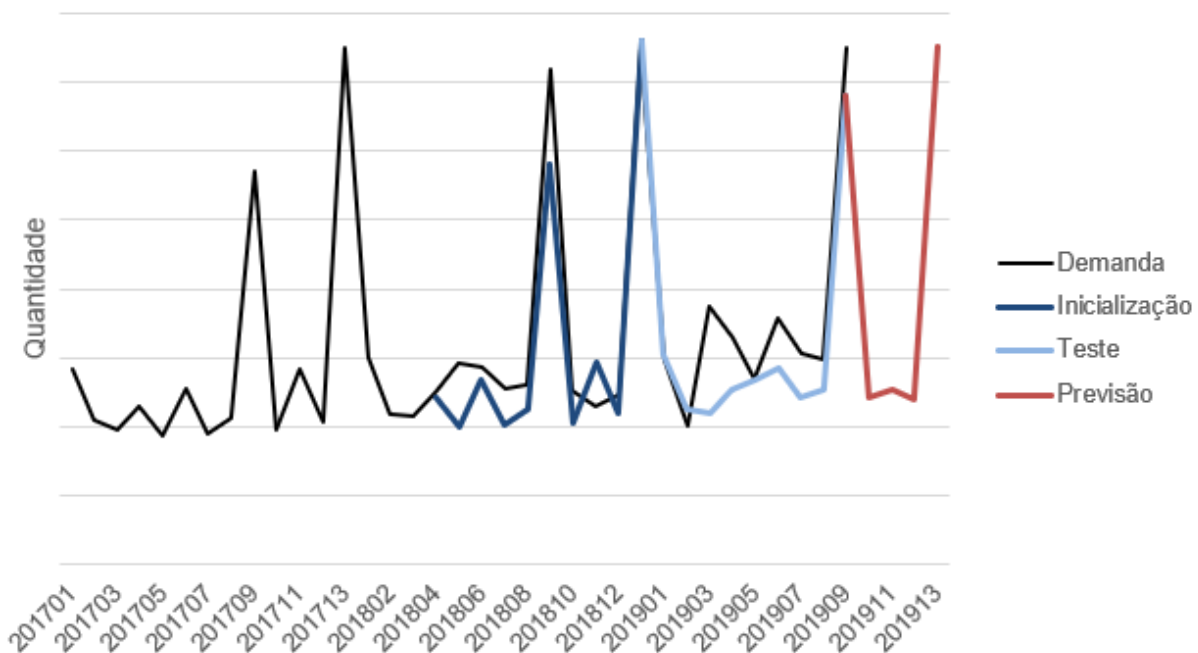
Fonte: elaborado pela autora.

### Cintos masculinos

Diferentemente da outra família de cintos, as previsões dos cintos masculinos obtiveram o menor erro quando utilizado o método de Holt Winter's com sazonalidade aditiva. Os coeficientes otimizados obtidos foram  $\alpha = 0,01$ ,  $\beta = 0,03$  e  $\gamma = 0,7$ .

A redução do MAPE, em relação ao modelo utilizado anteriormente, foi de 73% e o maior erro pontual, ocorrido na campanha 2019.03 foi de 41,4%.

Gráfico 32: Previsão da demanda de cintos masculinos



Fonte: elaborado pela autora.

Tabela 8: erros das previsões para cintos masculinos

	MAPE	MPE
<b>A</b>	17,9%	8,8%
<b>B</b>	17,4%	13,5%
<b>C</b>	40,0%	12,2%
<b>D</b>	15,8%	12,9%
<b>Antiga</b>	58,5%	

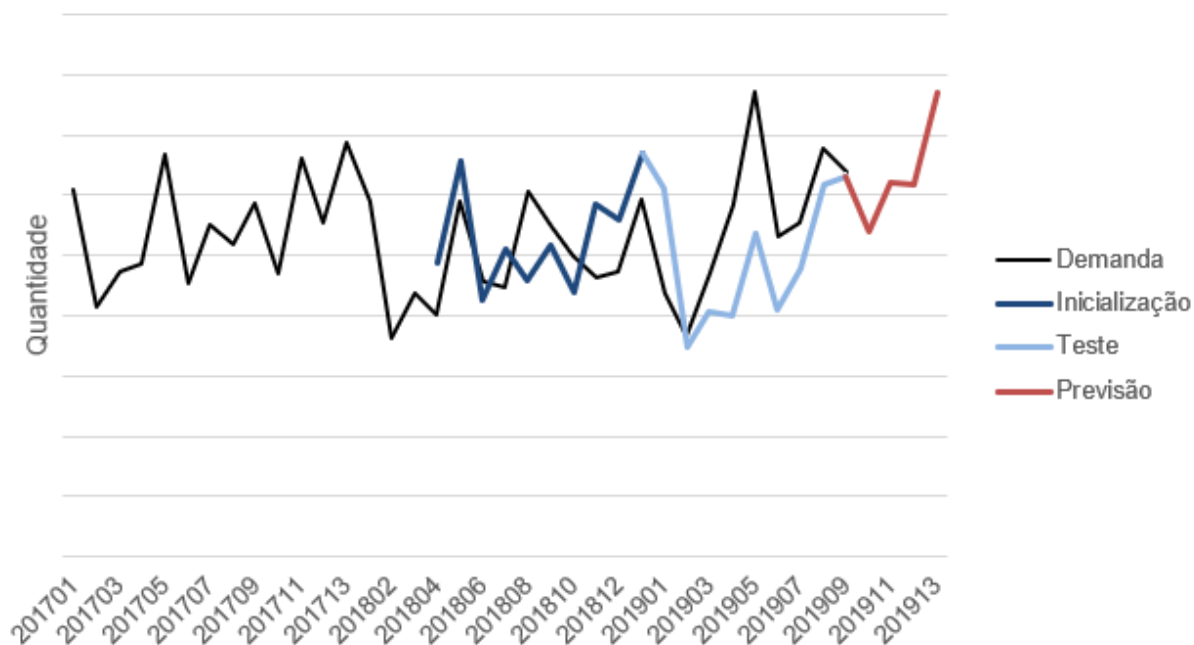
Fonte: elaborado pela autora.

### Carteiras femininas

A demanda das carteiras femininas foi mais bem prevista pelo método de suavização exponencial com sazonalidade multiplicativa. Os coeficientes usados para esses resultados foram  $\alpha = 0,44$  e  $\gamma = 0,7$ .

A redução do erro, em relação ao modelo anterior, foi de 68%. O maior erro pontual, da previsão da campanha 2019.01, foi de 39,4%.

Gráfico 33: Previsão da demanda de carteiras femininas



Fonte: elaborado pela autora.

Tabela 9: erros das previsões para carteiras femininas

	MAPE	MPE
<b>A</b>	18,5%	9,7%
<b>B</b>	20,6%	12,7%
<b>C</b>	22,8%	18,4%
<b>D</b>	41,4%	41,4%
<b>Antiga</b>	56,7%	

Fonte: elaborado pela autora.

## 5.9 Implementação do modelo

Uma vez testados os diversos métodos quantitativos, foi selecionado o que melhor se adequava a cada família conforme os resultados apresentados no item anterior. A opção de utilizar técnicas diferentes para cada uma delas não impacta o processo de forma negativa, uma vez que todos os modelos já haviam sido desenvolvidos para teste e não foi criada complexidade adicional para o time de planejamento que o utilizará. Pelo contrário, a utilização de métodos personalizados para cada família permite que sua demanda seja prevista conforme aquele que apresenta uma aderência maior à série. Dessa forma, cria-se uma metodologia de previsão quantitativa mais acurada para cada grupo, conforme suas características. O quadro 2 apresenta o resumo dos métodos selecionados para cada família, que devem ser utilizados com os coeficientes otimizados apresentados no item 5.8.

Quadro 2: métodos de previsão selecionados para cada família de produtos

Família de produtos	Método de previsão
<b>Bolsas neutras</b>	Holt Winter's com sazonalidade multiplicativa
<b>Bolsas coloridas</b>	Suavização com sazonalidade multiplicativa
<b>Cintos femininos</b>	Suavização com sazonalidade aditiva
<b>Cintos masculinos</b>	Holt Winter's com sazonalidade aditiva
<b>Carteiras femininas</b>	Suavização com sazonalidade multiplicativa

Fonte: elaborado pela autora.

A construção do modelo quantitativo e do processo de previsão foi realizada pelo time da Bain incorporando as sugestões e pedidos da gerência da empresa cliente e de seu time de planejamento comercial. Isso foi feito por meio de interações semanais nas quais o andamento

do projeto era apresentado pela consultoria e adaptado conforme os comentários e *feedback* dos envolvidos da Varejo SA.

Os resultados do método quantitativo, que reduziu os erros médios de previsão em quase 80%, foram muito bem recebidos pela gerência. O modelo foi apresentado em detalhe ao time de planejamento comercial, que também se mostrou muito satisfeito com sua simplicidade de operação e seus resultados; as funcionalidades e o passo a passo de como aplicá-lo foram minuciosamente explicados. As dúvidas e questionamentos do time foram sanadas o mais rápido possível para que o processo e modelo pudessem ser prontamente implementados no planejamento das campanhas a seguir. Como a conclusão do projeto ocorreu em meados de outubro, ao longo da campanha 11 de 2019, espera-se que o planejamento da segunda campanha de 2020 – realizado no início da campanha 12 de 2019 – seja realizado conforme o processo definido, utilizando o novo modelo de previsão.





## 6 CONCLUSÃO

A conclusão desse trabalho será dividida em três tópicos principais. Primeiramente, será realizada uma breve síntese dos problemas abordados e metodologias utilizadas. Em seguida, uma análise das limitações e melhorias a serem implementadas futuramente. Por fim, uma exposição de possíveis desdobramentos.

### 6.1 Síntese e resultados

O trabalho foi desenvolvido em uma empresa que comercializa bens de consumo em diversos canais, tendo iniciado as vendas no varejo recentemente. Os processos de planejamento comercial ainda são pouco robustos e foram identificados dois principais problemas a serem abordados com maior urgência: a definição do sortimento de produtos e a falta de acurácia da previsão de demanda usada para programação da produção.

A partir do levantamento de referências bibliográficas relevantes e análise da situação da rede de lojas da empresa, foi identificada a necessidade de se realizar uma segmentação das lojas. Os segmentos obtidos, com suas características específicas, foram utilizados como insumo para redefinição do sortimento, que varia de acordo com as necessidades das lojas e clientes.

Com as redefinições realizadas para cada segmento obtido, a densidade de exposição (número de SKUs expostos/m<sup>2</sup>) dos segmentos 1-AB, 1-BC, 2-AB e 2-BC devem ser reduzidas em mais de 20%, aproximando ou superando a densidade de concorrentes utilizadas como parâmetro. É esperada, ainda, uma redução média de 9% do portfólio comercializado (em número de SKUs diferentes) com redução de apenas 3% na receita bruta. Vale ressaltar que esse impacto deve ser compensado ou superado pela compra de produtos da mesma categoria daqueles excluídos, por conta da melhor experiência oferecida ao cliente com um sortimento direcionado.

No que diz respeito ao segundo problema identificado, o estudo das previsões de demanda utilizadas para o planejamento da produção apontou níveis de erros muito elevados, que levam a uma produção desregulada com a demanda real. Para solucionar esse problema, foram seguidos os vários passos de procedimentos propostos na literatura sobre o assunto, desde a coleta e análise de dados históricos de demanda até o estudo de métodos quantitativos adequados, sua aplicação e análise dos resultados obtidos. O método quantitativo proposto

apresentou uma redução média de 79,4% dos erros de previsão, mesmo antes da incorporação dos inputs qualitativos pelo time de marketing.

Os resultados dos métodos utilizados foram apresentados à gerência da empresa, que se mostrou extremamente satisfeita com o aumento da acurácia das previsões. O time de planejamento responsável pela previsão da demanda recebeu o novo processo de forma bastante positiva, sendo capaz de operar o modelo sem grandes dificuldades após as demonstrações e explicações necessárias. A implementação do novo procedimento foi iniciada enquanto esse trabalho era finalizado, de forma que os resultados não puderam ser analisados. Estima-se, com os níveis de acurácia obtidos, que a ruptura de estoque anual seja reduzida de mais de 2% para menos de 1% da receita bruta, ou seja, no máximo 1% da receita anual será perdida por conta da falta de itens no estoque central. Apesar da redução do valor parecer pouco relevante, vale lembrar que os níveis de ruptura alcançados antes do projeto só eram possíveis devido à obtenção de itens do estoque de outros canais da Varejo SA. No futuro, o crescimento da operação do varejo impossibilitará a realização dessa medida paliativa, o que faria com que os níveis de ruptura aumentassem significativamente caso o antigo processo de previsão fosse mantido.

## 6.2 Análise crítica

A principal limitação encontrada para o desenvolvimento desse trabalho foi relacionada a utilização de softwares especializados como *Forecast Pro*® cujas licenças poderiam custar até R\$20 mil e outros que, mesmo com menores custos ou livres, demandariam o treinamento do time de planejamento para sua correta utilização. A gerência da Varejo SA apresentou resistência a realizar os investimentos necessários e o time de planejamento comercial priorizou o uso do Microsoft Excel, programa com o qual já apresentava afinidade. Devido à intenção de se utilizar continuamente as ferramentas desenvolvidas, para atualização da segmentação e previsões futuras, optou-se por realizar o trabalho no software que o time ficasse mais confortável de operar. Essa limitação significou maior esforço de modelagem, principalmente para realização de análises exploratórias – tal qual a de autocorrelação das séries históricas – que poderiam ter sido simplificadas e feitas automaticamente por esses programas.

Em relação à segmentação da rede de lojas, uma possível crítica é relacionada à validação da clusterização realizada. O coeficiente de silhueta obtido na solução escolhida, apesar de ser o mais alto dentre os testes, não chegou ao nível desejado. Um valor a partir de 0,51 aponta uma estrutura aceitável, de acordo com a interpretação subjetiva de Kaufman e

Rousseeuw (1990), enquanto o encontrado foi de 0,44. Espera-se que com a agregação de novos pontos ao banco de dados, conforme a expansão esperada da rede de lojas, a clusterização torne-se mais robusta.

No que diz respeito ao modelo de previsão utilizado, a imaturidade da operação do varejo e consequente escassez de dados também configuram uma restrição importante. Devido ao curto tempo de funcionamento da rede, havia apenas dois anos completos de histórico de vendas para iniciar e calibrar o modelo, o que restringiu a escolha de métodos a testar além de reduzir a acurácia dos resultados. Os resultados para carteiras femininas, menos satisfatórios, poderiam ser revistos utilizando-se outros métodos mais ajustados às características da demanda da família, com o histórico de dados adequado.

### **6.3 Desdobramentos**

No futuro, dada a expansão da rede de lojas, poderiam ser testadas metodologias diferentes de clusterização assim como outros métodos quantitativos para segmentação da rede. No entanto, deve sempre ser comparada a quantidade de informação relevante à simplicidade de operação do modelo, que deve ser facilmente utilizado e atualizado pelo time de planejamento comercial e marketing conforme a rede de lojas se desenvolve.

Em relação ao modelo de previsão de demanda, no médio a longo prazo poderão ser testados novos métodos que exigem mais ciclos de histórico para inicialização e que podem trazer resultados ainda mais acurados, como os de correlação. Dado o curto histórico atual, apesar de se conhecer qualitativamente a influência de certos fatores na demanda, é impossível quantificar e comprovar as relações de forma robusta. Em um futuro mais distante, a agregação de novos dados pode também motivar testes de realização da previsão diretamente no nível SKU.

Com o desenrolar da operação do varejo, o próprio time de marketing também reunirá mais conhecimento e expertise a respeito da indústria e consumidores, aumentando a qualidade dos inputs qualitativos incluídos após a fase de previsão quantitativa apresentada.



## 7 REFERÊNCIAS BIBLIOGRÁFICAS

- ARNOLD, J. R. T.; CHAPMAN, S. N.; CLIVE, L. M. **Introduction to Materials Management**. 6. ed. New Jersey: Pearson Prentice Hal, 2008.
- BALDERSTON, F. E. Assortment Choice in Wholesale and Retail Marketing. **Journal of Marketing**, v. 21, n. 2, p. 175, 1956.
- BERMINGHAM, P.; HERNANDEZ, T.; CLARKE, I. Network planning and retail store segmentation: A spatial clustering approach. **International Journal of Applied Geospatial Research**, v. 4, n. 1, p. 67–79, 2013.
- BLYTHE, J. **Essentials of Marketing**. 3rd. ed. London: Pearson Education, 2005.
- DOLNIČAR, S. Using cluster analysis for market segmentation - typical misconceptions, established methodological weaknesses and some recommendations for improvement. **Australasian Journal of Market Research**, v. 11, n. 2, p. 5–12, 2003.
- HART, C.; RAFIQ, M. The dimensions of assortment: A proposed hierarchy of assortment decision making. **International Review of Retail, Distribution and Consumer Research**, v. 16, n. 3, p. 333–351, 2006.
- HOLLANDER, S.C. Notes on the retail accordion. **Journal of Retailing**, New York, v.42, p. 29-40, 1966.
- KAUFMAN, L.; ROUSSEEUW, P. J. **Finding groups in data: An introduction to cluster analysis**. Hoboken: John Wiley and Sons, 1990b.
- KOTLER, P.; ARMSTRONG, G. **Principles of Marketing**. 16. ed. [s.l.] Pearson Education, 2016.
- LEE, S.E.; KUNZ, G.I. Assortment diversity in relation to financial productivity: contribuTions towards merchandising theory. **The Journal of Fashion Marketing and Management**, New York, v.5, p. 303–312, 2001.
- MAKRIDAKIS, S.; WHEELWRIGHT, S. C.; MCGEE, V. E. **Forecasting: Methods and Applications**. 3. ed. New York: John Wiley and Sons, 1998.
- MOUTINHO, L. Quantitative Analysis in Marketing. In: **The Marketing Book**. 6. ed. [s.l.] Elsevier, 2003. p. 170–220.
- ROUSSEEUW, P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. **Journal of Computational and Applied Mathematics**, v. 20, n. C, p. 53–65, 1987.
- SAUNDERS, J. A. Cluster analysis for market segmentation. **European Journal of Marketing**, v. 14, n. 7, p. 422–435, 1980.

SIMONSON, I. The effect of product assortment on buyer preferences. **Journal of Retailing**, New York, v.75, p. 347-370, 1999.

TAN, P.-N.; STEINBACH, M.; KUMAR, V. **Introduction to Data Mining**. New York: Pearson Education, 2006.

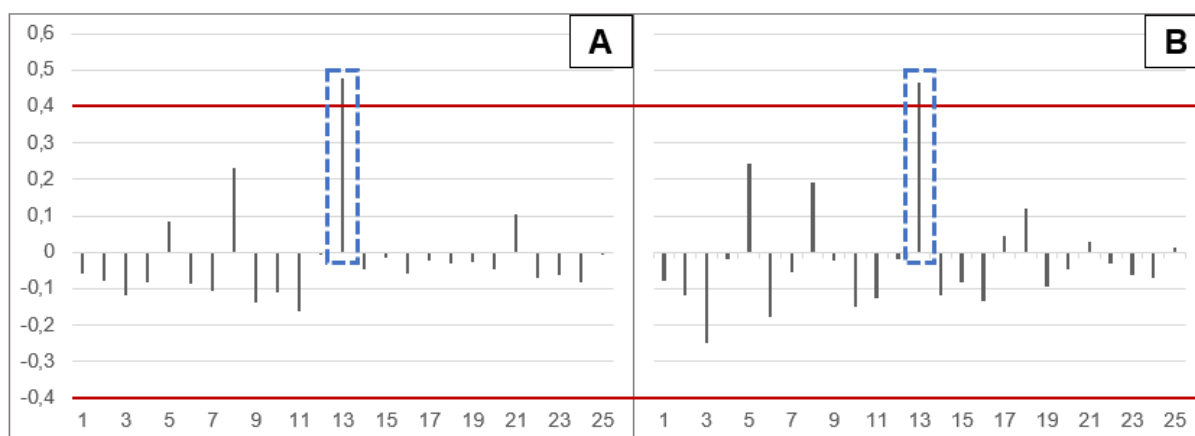
## ANEXO A: AUTOCORRELAÇÃO DAS SÉRIES HISTÓRICAS

A análise da autocorrelação das demandas históricas das famílias estudadas contribuiu para a comprovação da existência – ou ausência – de sazonalidade. Conforme levantado no capítulo 3, um índice de autocorrelação  $r_k$  significativo, que supera os limites calculados, é indício de forte relação entre demandas defasadas de  $k$  períodos, indicando uma sazonalidade dessa amplitude.

Foram calculados os índices de autocorrelação para as cinco famílias estudadas nesse trabalho, com base na demanda histórica por campanha de dois anos completos (2017 e 2018), totalizando 26 períodos. É importante lembrar que, no caso da Varejo SA, a previsão é realizada por campanha e não por mês, logo uma sazonalidade anual existirá quando  $r_{13}$  for significativo.

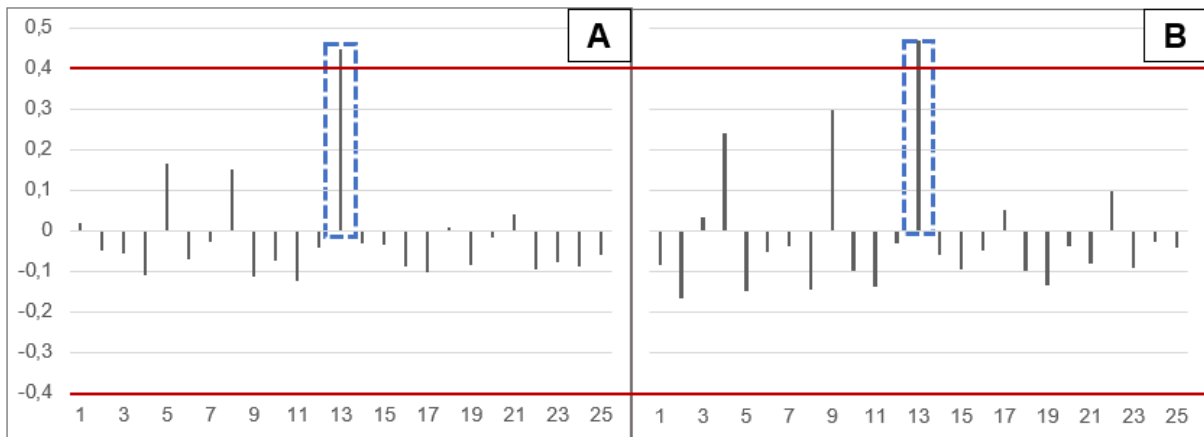
Dessa forma, foram identificadas claras sazonalidades anuais tanto pelos gráficos de sazonalidade quanto pela autocorrelação para as quatro famílias de bolsas e cintos, enquanto as carteiras femininas não apresentaram essa característica clara.

Gráfico 34: correlograma da demanda de bolsas neutras (A) e coloridas (B)



Fonte: elaborado pela autora a partir de dados da Varejo SA

Gráfico 35: correlograma da demanda de cintos femininos (A) e masculinos (B)



Fonte: elaborado pela autora a partir de dados da Varejo SA

Gráfico 36: correlograma da demanda de carteiras femininas



Fonte: elaborado pela autora a partir de dados da Varejo SA