

**PALOMA AYUMI ISHII**

# **Estudo do comportamento das vendas em uma empresa de sorvetes**

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

**São Paulo**

**2005**

**PALOMA AYUMI ISHII**

# **Estudo do comportamento das vendas em uma empresa de sorvetes**

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

**Orientador: Prof. Dr. Alberto W. Ramos**

**São Paulo**

**2005**

## **Agradecimentos**

Gostaria de dedicar este espaço para agradecer a todos aqueles que, direta ou indiretamente, colaboraram para a conclusão deste trabalho.

Em primeiro lugar, o meu muito obrigada a minha família, especialmente a meus pais, Masao e Ana, por terem me proporcionado a educação e a base familiar necessária para que eu pudesse chegar até aqui.

A toda equipe de Planejamento da Kibon, principalmente à Flavia Sebastiani, Paula Schreiber e Gustavo Benitez, pela confiança depositada e por todo aprendizado profissional que adquiri no ano de estágio inesquecível que tive na Kibon.

Ao professor Alberto W. Ramos pela valiosa orientação na condução deste trabalho.

Por fim, não poderia deixar de agradecer às minhas amigas queridas e a todos aqueles que estiveram ao meu lado neste ano, por todo incentivo e compreensão.

## **Resumo**

O presente trabalho tem como objetivo a análise das vendas de uma empresa de sorvetes, através da aplicação de métodos estatísticos e modelos de previsão de demanda. A temperatura das regiões onde se localizam as filiais da empresa é apresentada como sendo um possível fator que exerce impacto significativo sobre as vendas; essa relação é verificada através de uma análise de correlação entre variáveis e é testado um modelo de Regressão Linear na tentativa de prever as demandas futuras. Além disso, alguns modelos ARIMA são aplicados aos dados para verificar a influência das vendas de períodos anteriores sobre as vendas reais verificadas. Por fim, através dos resultados obtidos são sugeridos alguns trabalhos futuros, como a estruturação do procedimento de coleta de dados e a possibilidade de uso de outros métodos para uma análise mais aprofundada do comportamento das vendas da empresa e dos fatores que exerceriam impacto sobre as mesmas.

## **Abstract**

This study has as objective the analysis of the sales of an Ice Cream Company, through the application of statistical and demand forecast methods. The temperature of the regions where the company's distribution centers are located is presented as being a possible factor that exerts significant impact on the sales volume; this assumption is verified through a correlation analysis between the variables and a linear regression model is tested in the attempt to foresee the demand. Moreover, some ARIMA models are applied to the data to verify the influence of previous periods sales on the current ones. Finally, through the analysis of the achieved results some future works are suggested, such as the definition of a data capture procedure and the use of some other methods for a deeper analysis of the company's sales behavior and the factors that would exert impact on them.

# SUMÁRIO

|                     |                                                            |                  |
|---------------------|------------------------------------------------------------|------------------|
| <b><u>1.</u></b>    | <b><u>INTRODUÇÃO</u></b>                                   | <b><u>2</u></b>  |
| <b><u>2.</u></b>    | <b><u>A EMPRESA</u></b>                                    | <b><u>5</u></b>  |
| <b><u>2.1</u></b>   | <b><u>HISTÓRICO</u></b>                                    | <b><u>5</u></b>  |
| <b><u>2.2</u></b>   | <b><u>ESTRUTURA DA EMPRESA</u></b>                         | <b><u>8</u></b>  |
| <b><u>2.3</u></b>   | <b><u>MARCAS E PRODUTOS</u></b>                            | <b><u>11</u></b> |
| <b><u>2.4</u></b>   | <b><u>ÁREAS DA EMPRESA</u></b>                             | <b><u>13</u></b> |
| <b><u>3.</u></b>    | <b><u>O PROCESSO DE PLANEJAMENTO</u></b>                   | <b><u>15</u></b> |
| <b><u>3.1</u></b>   | <b><u>Visão Geral</u></b>                                  | <b><u>15</u></b> |
| <b><u>3.2</u></b>   | <b><u>PLANEJAMENTO DA DEMANDA</u></b>                      | <b><u>17</u></b> |
| <b><u>3.2.1</u></b> | <b><u>COMPONENTES DA PREVISÃO</u></b>                      | <b><u>17</u></b> |
| <b><u>3.2.2</u></b> | <b><u>PROCESSO DE PLANEJAMENTO DA DEMANDA</u></b>          | <b><u>18</u></b> |
| <b><u>4.</u></b>    | <b><u>O PROBLEMA E OBJETIVO</u></b>                        | <b><u>21</u></b> |
| <b><u>4.1</u></b>   | <b><u>Visão Geral</u></b>                                  | <b><u>21</u></b> |
| <b><u>4.2</u></b>   | <b><u>DEFINIÇÃO DO PROBLEMA</u></b>                        | <b><u>23</u></b> |
| <b><u>4.3</u></b>   | <b><u>OBJETIVOS DO TRABALHO</u></b>                        | <b><u>24</u></b> |
| <b><u>5.</u></b>    | <b><u>REVISÃO BIBLIOGRÁFICA</u></b>                        | <b><u>26</u></b> |
| <b><u>5.1</u></b>   | <b><u>MODELOS DE PREVISÃO</u></b>                          | <b><u>26</u></b> |
| <b><u>5.2</u></b>   | <b><u>ETAPAS DE APLICAÇÃO DE UM MODELO DE PREVISÃO</u></b> | <b><u>28</u></b> |
| <b><u>5.3</u></b>   | <b><u>DECOMPOSIÇÃO DE SÉRIES TEMPORAIS</u></b>             | <b><u>31</u></b> |
| <b><u>5.4</u></b>   | <b><u>MÉDIAS</u></b>                                       | <b><u>32</u></b> |
| <b><u>5.4.1</u></b> | <b><u>MÉDIA SIMPLES</u></b>                                | <b><u>32</u></b> |
| <b><u>5.4.2</u></b> | <b><u>MÉDIA MÓVEL</u></b>                                  | <b><u>32</u></b> |
| <b><u>5.5</u></b>   | <b><u>REGRESSÃO LINEAR</u></b>                             | <b><u>33</u></b> |
| <b><u>5.5.1</u></b> | <b><u>COEFICIENTE DE CORRELAÇÃO</u></b>                    | <b><u>33</u></b> |
| <b><u>5.5.2</u></b> | <b><u>REGRESSÃO LINEAR SIMPLES</u></b>                     | <b><u>35</u></b> |
| <b><u>5.5.3</u></b> | <b><u>REGRESSÃO LINEAR MÚLTIPLA</u></b>                    | <b><u>38</u></b> |
| <b><u>5.6</u></b>   | <b><u>ARIMA</u></b>                                        | <b><u>42</u></b> |
| <b><u>6.</u></b>    | <b><u>DEFINIÇÃO E LEVANTAMENTO DOS DADOS</u></b>           | <b><u>49</u></b> |
| <b><u>6.1</u></b>   | <b><u>CONSIDERAÇÕES INICIAIS</u></b>                       | <b><u>49</u></b> |
| <b><u>6.2</u></b>   | <b><u>OBJETO DE ESTUDO E PREVISÃO</u></b>                  | <b><u>49</u></b> |
| <b><u>6.2.1</u></b> | <b><u>OBJETO DE PREVISÃO</u></b>                           | <b><u>49</u></b> |
| <b><u>6.2.2</u></b> | <b><u>FATORES A SEREM ESTUDADOS</u></b>                    | <b><u>50</u></b> |
| <b><u>6.2.3</u></b> | <b><u>PRODUTOS ANALISADOS</u></b>                          | <b><u>52</u></b> |
| <b><u>6.2.4</u></b> | <b><u>FILIAIS ESCOLHIDAS</u></b>                           | <b><u>55</u></b> |
| <b><u>6.3</u></b>   | <b><u>PERÍODO DE TEMPO</u></b>                             | <b><u>56</u></b> |

|                     |                                                 |            |
|---------------------|-------------------------------------------------|------------|
| <b><u>7.</u></b>    | <b><u>ANÁLISE</u></b>                           | <b>58</b>  |
| <b><u>7.1</u></b>   | <b><u>TEMPERATURA (T) X VOLUME (T)</u></b>      | <b>58</b>  |
| <b><u>7.2</u></b>   | <b><u>TEMPERATURA (T-1) X VOLUME (T)</u></b>    | <b>61</b>  |
| <b><u>7.3</u></b>   | <b><u>REGRESSÃO LINEAR</u></b>                  | <b>64</b>  |
| <b><u>7.4</u></b>   | <b><u>AUTOCORRELAÇÃO</u></b>                    | <b>66</b>  |
| <b><u>7.5</u></b>   | <b><u>MODELO ARIMA</u></b>                      | <b>69</b>  |
| <b><u>8.</u></b>    | <b><u>RESULTADOS</u></b>                        | <b>82</b>  |
| <b><u>9.</u></b>    | <b><u>TRABALHOS FUTUROS</u></b>                 | <b>85</b>  |
| <b><u>9.1</u></b>   | <b><u>PROCEDIMENTO PARA COLETA DE DADOS</u></b> | <b>85</b>  |
| <b><u>9.1.1</u></b> | <b><u>FATORES E ATRIBUTOS</u></b>               | <b>86</b>  |
| <b><u>9.2</u></b>   | <b><u>ANÁLISES FUTURAS</u></b>                  | <b>90</b>  |
| <b><u>10.</u></b>   | <b><u>CONCLUSÃO</u></b>                         | <b>93</b>  |
| <b><u>11.</u></b>   | <b><u>ANEXOS</u></b>                            | <b>96</b>  |
| <b><u>12.</u></b>   | <b><u>REFERÊNCIA BIBLIOGRÁFICA</u></b>          | <b>104</b> |

## ÍNDICE DE FIGURAS

|                                                                                                        |    |
|--------------------------------------------------------------------------------------------------------|----|
| <a href="#"><u>Figura 2.1 - Cronologia da Unilever</u></a>                                             | 7  |
| <a href="#"><u>Figura 2.2 - Principais marcas de HPC</u></a>                                           | 8  |
| <a href="#"><u>Figura 2.3 - Principais marcas da Unilever Foods</u></a>                                | 8  |
| <a href="#"><u>Figura 2.4 – Fábricas e escritórios da Unilever no Brasil</u></a>                       | 9  |
| <a href="#"><u>Figura 2.5 - Localização das Filiais da Kibon</u></a>                                   | 10 |
| <a href="#"><u>Figura 2.6 - Distribuição das vendas de 2004 por categoria</u></a>                      | 12 |
| <a href="#"><u>Figura 2.7 - Organograma da empresa e, em destaque, a área de Planejamento</u></a>      | 13 |
| <a href="#"><u>Figura 3.1 - Fluxograma do processo de Planejamento</u></a>                             | 16 |
| <a href="#"><u>Figura 3.2 - Componentes da Previsão de demanda</u></a>                                 | 17 |
| <a href="#"><u>Figura 3.3 - Processo de Planejamento da Demanda</u></a>                                | 19 |
| <a href="#"><u>Figura 4.1 - Comportamento do Consumidor Brasileiro</u></a>                             | 21 |
| <a href="#"><u>Figura 5.1 - Etapas para escolha e aplicação dos modelos de previsão</u></a>            | 30 |
| <a href="#"><u>Figura 5.2 - Modelo para identificação do modelo ARIMA adequado</u></a>                 | 46 |
| <a href="#"><u>Figura 6.1 - Classificação ABC das famílias de produtos Kibon</u></a>                   | 53 |
| <a href="#"><u>Figura 6.2 - Classificação ABC das famílias da linha Impulse</u></a>                    | 54 |
| <a href="#"><u>Figura 6.3 - Classificação ABC das famílias da linha Take Home</u></a>                  | 54 |
| <a href="#"><u>Figura 7.1 - Gráfico Temperatura x Vendas (São Paulo)</u></a>                           | 60 |
| <a href="#"><u>Figura 7.2 - Gráfico Temperatura x Vendas (Recife)</u></a>                              | 60 |
| <a href="#"><u>Figura 7.3 - Autocorrelação de C (Filial São Paulo)</u></a>                             | 67 |
| <a href="#"><u>Figura 7.4 - Autocorrelação Parcial de C (Filial São Paulo)</u></a>                     | 67 |
| <a href="#"><u>Figura 7.5 - Autocorrelação de C (Filial Recife)</u></a>                                | 67 |
| <a href="#"><u>Figura 7.6 - Autocorrelação Parcial de C (Filial Recife)</u></a>                        | 68 |
| <a href="#"><u>Figura 7.7 - Função de autocorrelação até a ordem 52</u></a>                            | 69 |
| <a href="#"><u>Figura 7.8 - Vendas da família A na Filial São Paulo</u></a>                            | 70 |
| <a href="#"><u>Figura 7.9 - Comparação entre o teste normal dos modelos (4,0,3) e (4,0,0)</u></a>      | 72 |
| <a href="#"><u>Figura 7.10 - Função FAC do modelo (4,0,3) para A na Filial Recife</u></a>              | 72 |
| <a href="#"><u>Figura 7.11 - Função FACP do modelo (4,0,3) para A na Filial Recife</u></a>             | 73 |
| <a href="#"><u>Figura 7.12 - FAC para Família E (Filial Recife)</u></a>                                | 74 |
| <a href="#"><u>Figura 7.13 - FACP para Família E (Filial Recife)</u></a>                               | 74 |
| <a href="#"><u>Figura 7.14 - Função FACP dos resíduos da Família B (Filial Rio de Janeiro)</u></a>     | 77 |
| <a href="#"><u>Figura 7.15 - Gráfico de probabilidade normal para Família A (Filial Fortaleza)</u></a> | 78 |



|                                                                                                              |    |
|--------------------------------------------------------------------------------------------------------------|----|
| <a href="#"><u>Figura 7.16 - Gráfico de probabilidade normal para Família C (Filial São Paulo)</u></a> ..... | 78 |
| <a href="#"><u>Figura 7.17 - Gráfico de resíduos da Família A (Filial Rio de Janeiro)</u></a> .....          | 79 |
| <a href="#"><u>Figura 7.18 - Gráfico de resíduos da Família D (Filial Recife)</u></a> .....                  | 80 |
| <a href="#"><u>Figura 9.1 - Tabelas para coleta de dados</u></a> .....                                       | 88 |
| <a href="#"><u>Figura 9.2 - Representação de uma série qualquer com variável binária</u></a> .....           | 91 |

## ÍNDICE DE TABELAS

|                                                                                                           |    |
|-----------------------------------------------------------------------------------------------------------|----|
| <a href="#"><u>Tabela 2.1 - Características das áreas de vendas</u></a> .....                             | 10 |
| <a href="#"><u>Tabela 2.2 - Divisão dos produtos da empresa</u></a> .....                                 | 12 |
| <a href="#"><u>Tabela 5.1 - Métodos de previsão Quantitativos</u></a> .....                               | 27 |
| <a href="#"><u>Tabela 5.2 - FAC e FACP para identificação do modelo ARIMA</u></a> .....                   | 46 |
| <a href="#"><u>Tabela 7.1 - Coeficientes de correlação entre Volume e Temperatura</u></a> .....           | 59 |
| <a href="#"><u>Tabela 7.2 - <math>R^2</math> entre Volume e Temperatura</u></a> .....                     | 59 |
| <a href="#"><u>Tabela 7.3 - Valor P da correlação Volume-Temperatura</u></a> .....                        | 59 |
| <a href="#"><u>Tabela 7.4 - Coeficientes de correlação entre Volume (t) e Temperatura (t-1)</u></a> ..... | 61 |
| <a href="#"><u>Tabela 7.5 - <math>R^2</math> entre Volume (t) e Temperatura (t-1)</u></a> .....           | 61 |
| <a href="#"><u>Tabela 7.6 - Valor P da correlação Volume (t)-Temperatura (t-1)</u></a> .....              | 61 |
| <a href="#"><u>Tabela 7.7 - Maiores <math>R^2</math> e respectivos Valores P</u></a> .....                | 62 |
| <a href="#"><u>Tabela 7.8 - Diferença entre os <math>R^2</math></u></a> .....                             | 62 |
| <a href="#"><u>Tabela 7.9 - Resultados da Regressão Linear (Filial São Paulo)</u></a> .....               | 65 |
| <a href="#"><u>Tabela 7.10 - Resultados da Regressão Linear (Filial Rio de Janeiro)</u></a> .....         | 65 |
| <a href="#"><u>Tabela 7.11 - Autocorrelação Parcial de ordem 1 por Filial e Família</u></a> .....         | 68 |
| <a href="#"><u>Tabela 7.12 - Modelos ARIMA escolhidos</u></a> .....                                       | 75 |
| <a href="#"><u>Tabela 7.13 – Erro Absoluto Percentual Médio (ARIMA)</u></a> .....                         | 77 |
| <a href="#"><u>Tabela 9.1 - Resumos dos fatores a serem levantados</u></a> .....                          | 89 |

## **LISTA DE SIGLAS**

**ARIMA:** Autoregressive Integrated Moving Average

**AP:** Annual Plan

**CD:** Centro de Distribuição

**FAC:** Função de Autocorrelação

**FACP:** Função de Autocorrelação Parcial

**HPC:** Home and Personal Care

**IC:** Ice Dream

**MPS:** Master Planning Schedule

**OP:** Operational Plan

**SFE:** Sales Forecast Error

**SKU:** Stock Keeping Unit

**SOA:** Share of Assortment

## **Introdução**

A thick, horizontal orange bar with a slight gradient and a wavy, hand-drawn appearance, extending across the lower portion of the page.

## 1. INTRODUÇÃO

Este trabalho tem como proposta estudar os fatores que exerceriam impacto sobre as vendas da Kibon, utilizando métodos e modelos estatísticos. Dessa maneira, espera-se possibilitar à empresa um maior conhecimento sobre suas ações e também sobre as consequências da influência de fatores externos, podendo melhorar a previsão de demanda da empresa.

No capítulo 2, serão apresentados um breve histórico da empresa e uma descrição de seus produtos e áreas.

Em seguida, será detalhado o processo de Planejamento como um todo e as respectivas partes que o compõem: o planejamento da produção, o de materiais e a programação da produção. Uma seção é dedicada especialmente para o planejamento de demanda, detalhando como se dá o envolvimento das diversas áreas da empresa e o fluxo de informações neste processo. O método de previsão utilizado pela empresa é definido e são descritos também todos os produtos desta etapa de planejamento da demanda.

O capítulo 4 é dedicado ao problema a ser tratado neste trabalho e o objetivo do mesmo. Inicialmente, é feita uma breve descrição do contexto de mercado onde a empresa está inserida, assim como os impactos causados pela falta do produto nos pontos de venda, para em seguida serem definidos o problema e o objetivo do estudo.

O capítulo 5 refere-se à revisão bibliográfica, iniciando pelas classificações dos modelos de previsão. Em seguida, são apresentados os principais modelos de previsão de demanda, entre eles o de Regressão Linear e o modelo ARIMA. Outros tipos de modelo podem ser encontrados também nos anexos.

Antes de iniciar a análise em si, os objetos de estudo, no caso os fatores que exerceriam impacto sobre as vendas da Kibon, são definidos e é feito o levantamento de dados no capítulo 6. Nele é explicado o porquê da escolha de certos fatores, produtos e filiais para esta análise.

O capítulo 7 trata dos resultados das análises feitas, relacionando o fator escolhido no capítulo anterior com o volume de vendas da empresa, assim como são testados alguns modelos de previsão. Além disso, uma análise de autocorrelação é feita para verificar se há relação entre as vendas de duas ou mais semanas consecutivas.

Os resultados obtidos nesta análise seguem no capítulo 8.

Com base nesses resultados, levantou-se a necessidade e possibilidade de trabalhos futuros, contidos no capítulo 9.

Por fim, as dificuldades encontradas para a realização deste trabalho, assim como os aprendizados obtidos são apresentados no capítulo de conclusão.

**Empresa**

## **2. A EMPRESA**

### **2.1 Histórico**

- **Unilever**

A Unilever, multinacional de origem anglo-holandesa, chegou ao Brasil em 1929, através da fundação da Sociedade Anônima Irmãos Lever, que iniciou as suas atividades comercializando o sabão Sunlight, primeiro produto criado na Inglaterra por William Hesketh Lever, um dos fundadores da empresa, em 1890.

Em 1930, começou a funcionar na Vila Anastácio, em São Paulo, a primeira fábrica da companhia, que deixou de importar o sabão Sunlight e passou a produzi-lo no Brasil.

Em 1944, a Irmão Lever iniciou o processo de compras de outras empresas através da compra da Perfumaria J. & E. Atkinsons; no entanto, sua primeira grande e significativa aquisição foi a da Companhia Gessy Industrial, em 1960, dando origem assim às Indústrias Gessy Lever.

A entrada da empresa no setor de alimentos ocorreu em 1970, com o surgimento da divisão Van der Bergh e a comercialização no país da primeira margarina cremosa, a Doriana. Posteriormente, a participação no mercado desse produto cresceu ainda mais com a aquisição da Anderson Clayton.

No entanto, a adoção do nome internacional Unilever só ocorreu em 2000, após a aquisição mundial da Bestfoods, que no Brasil era conhecida como Refinações de Milho Brasil (RMB) e era dona das marcas Maisena, Hellmann's, Knorr e Arisco, entre outras.



- **Kibon**

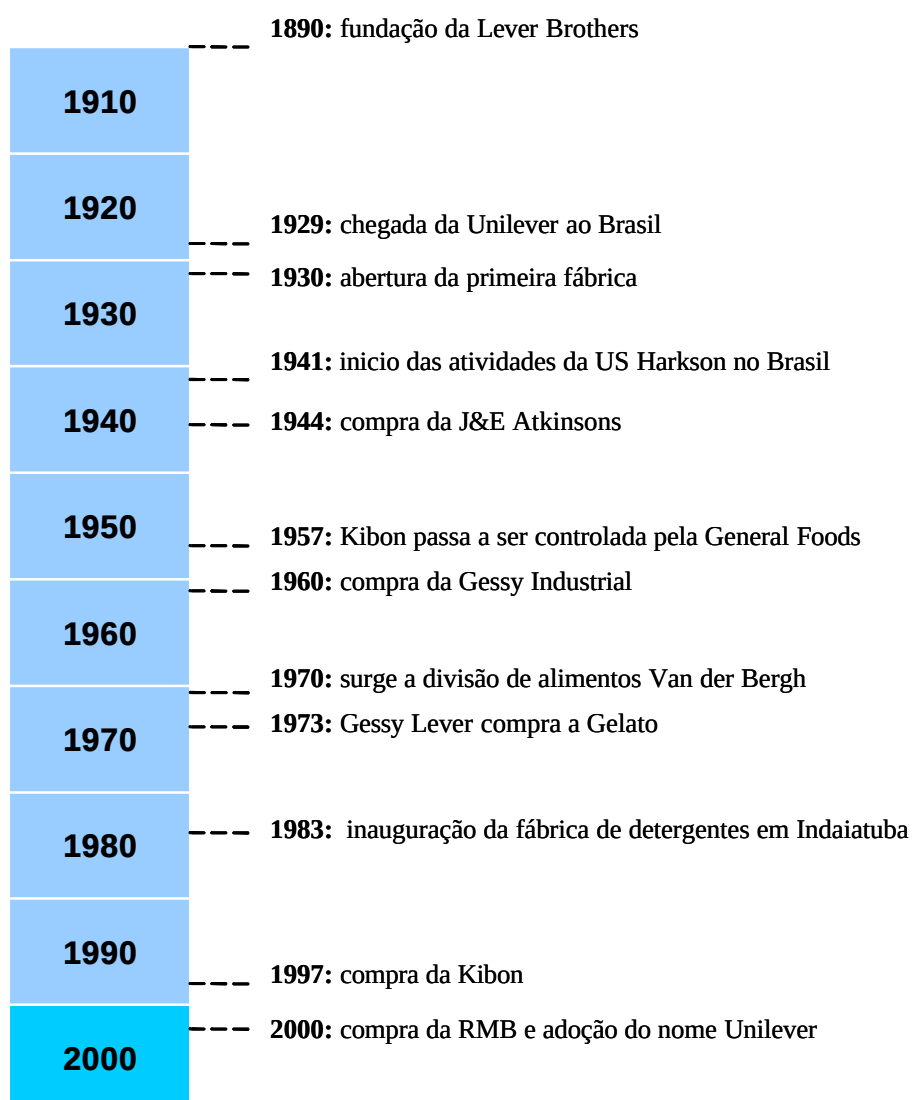
A fabricação de sorvetes em escala industrial no Brasil teve início em 1941, quando a americana Companhia U.S. Harkson do Brasil começou a funcionar no Rio de Janeiro. Somente dez anos mais tarde, foi acrescentado ao nome da companhia o “Kibon”.

Os primeiros produtos foram o Chicabon e o Eskibon, que logo se tornaram um sucesso e que são fabricados até hoje, sendo um dos picolés mais vendidos pela empresa.

Em 1957, a empresa passou a ser controlada pela General Foods Corporation e dominou o mercado, praticamente sem nenhum concorrente, até o começo da década de 70, quando a marca Gelato surgiu no país. Dois anos mais tarde, a Gelato foi comprada pela Gessy Lever.

Na mesma época, a Nestlé entrou no ramo de sorvetes com o lançamento da Yopa e, em 1990, a Gessy Lever e a Nestlé formaram uma joint venture para abrigar a Gelato e a Yopa, sociedade que só foi desfeita em 1994.

Finalmente, em 1997, a Gessy Lever voltou ao mercado de sorvetes com a compra da Kibon, que se caracterizou como sendo a maior aquisição da Unilever fora dos Estados Unidos em sessenta anos.



*Figura 2.1 - Cronologia da Unilever*

*Elaborado pela autora*

## 2.2 Estrutura da empresa

- **Divisões**

A Unilever no Brasil está dividida em quatro divisões. São elas:

- **HPC (Home and Personal Care):** corresponde à divisão que cuida dos produtos de higiene e beleza. A empresa é líder de mercado em diversos segmentos nos quais atua através dos produtos desta divisão, cujas principais marcas estão na figura 2.2.



Figura 2.2 - Principais marcas de HPC

Fonte: Empresa

- **Unilever Foods:** corresponde aos produtos alimentícios.



Figura 2.3 - Principais marcas da Unilever Foods

Fonte: Empresa

- **Unilever Food Solutions:** corresponde aos produtos voltados para cadeias de alimentação, como restaurantes.

➤ **IC (Ice Cream – Kibon):** divisão de sorvetes, líder no mercado brasileiro. Será nesta divisão que o presente trabalho será desenvolvido.

- **Fábricas e escritórios**

A Unilever possui hoje no Brasil 13 fábricas localizadas nos estados de São Paulo, Minas Gerais, Goiás e Pernambuco, sendo que duas delas são da Kibon: uma na cidade de São Paulo e outra em Recife.

Já os escritórios concentram-se em São Paulo, em três principais locais: na Vila Anastácio, no Centro Empresarial e no Brooklin.



*Figura 2.4 – Fábricas e escritórios da Unilever no Brasil*

*Fonte: Empresa*

- **Estrutura de Vendas – Kibon**

Atualmente, a estrutura de vendas da Kibon está dividida em duas áreas: Área Direta e Indireta.

As principais características de cada uma das áreas estão resumidas na tabela 2.1.

A figura 2.5 apresenta a localização das filiais da Kibon.

Tabela 2.1 - Características das áreas de vendas

|                      | ÁREA DIRETA                   | ÁREA INDIRETA      |
|----------------------|-------------------------------|--------------------|
| Atendimento          | via Filiais da Kibon          | via Distribuidores |
| Representatividade   | 50% do volume                 | 50% do volume      |
| Cobertura geográfica | 10% do Brasil                 | 90% do Brasil      |
| Faturamento          | Maior (maior margem de lucro) | Menor              |

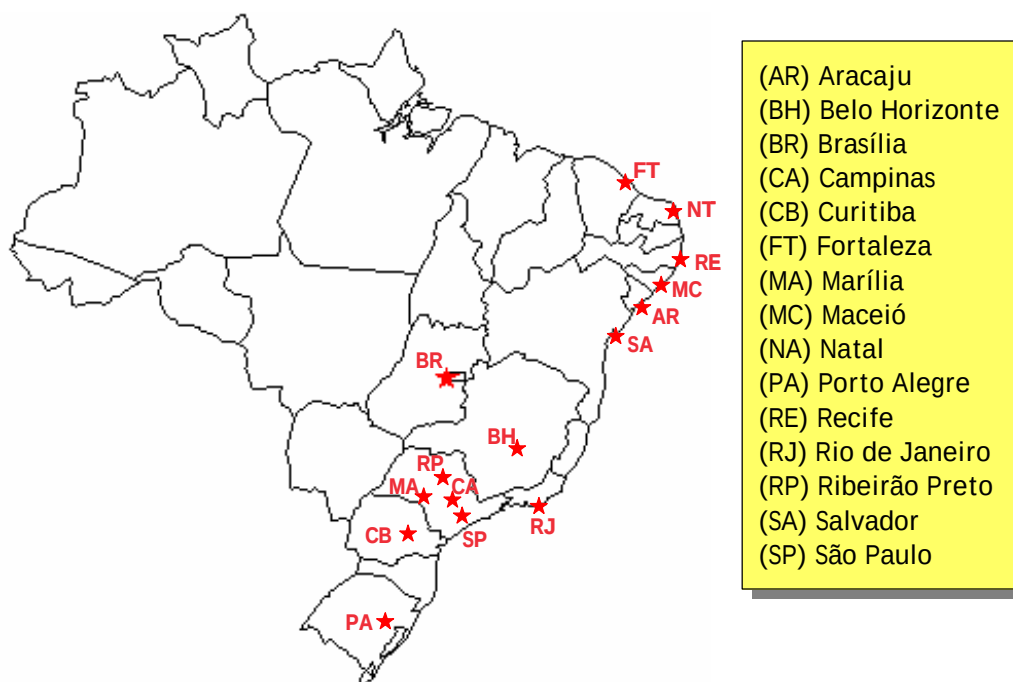


Figura 2.5 - Localização das Filiais da Kibon

Fonte: Empresa

- **Funcionários**

No mundo, a Unilever possui mais de 230.000 funcionários, sendo que 13.000 deles estão no Brasil.

Na Kibon, esse número de funcionários varia bastante devido às contratações nas temporadas de verão, mas o número médio seria de 1.560 pessoas.

## 2.3 Marcas e Produtos

Atualmente, a Kibon adota a seguinte divisão em seu portfólio de produtos: ***In home*** e ***Out of Home***.

Os produtos *In Home* correspondem àqueles que são consumidos “internamente”, nos lares das famílias, enquanto os de *Out of Home* têm como característica principal o fato de serem consumidos “externamente”, nas ruas, e ainda podem ser divididos em outros dois subgrupos: *Impulse* e *Food Service*.

A Kibon é dona de algumas das mais fortes marcas do mercado de sorvetes, possuindo atualmente uma variedade de mais de 50 produtos diferentes, entre marcas e sabores. São elas:

- **Cornetto:** Crocante, Brigadeiro, Royal Morango, Royal Caramelo, Copo Frutas Vermelhas, Copo Sundae e linha Aphrodiziac (Me dei bem, Uh Baby e Vai rolar).
- **Magnum:** Clássico, Special e Amêndoas.
- **Fruttare:** Uva, Limão, Abacaxi, Coco e Banana.
- **Clássicos:** Eskibon, Mini Eskibon, Chicabon, Tablito e Brigadeiro.
- **Ades:** Maça e Pêssego.
- **Kids:** Dedo Mola, Patrick, Yu Gi Oh, Stika Puxa, Chicabonzinho, Frutilly Morango, Frutilly Floc e Pop Cor.
- **Carte Dor:** Morando Merengue, Caramelo Toffee, Torta Creme de Maracujá, Creme Crocante, Limão, Pavê, Papaya, Trufa e Banana.
- **Pote 1 litro light:** morango e creme.
- **Potes 2 litros:** Chicabon, Creme, Flocos, Napolitano, Passas ao Rum, Morango, Cajá e Frutas.
- **Potes 5 litros:** Flocos, Chicabon, Brigadeiro, Abacaxi, Chocolate, Morango, Creme, Passas ao Rum, Crocante, Napolitano e Coco.

Tabela 2.2 - Divisão dos produtos da empresa

| Linha   | IN HOME        | OUT OF HOME |                |
|---------|----------------|-------------|----------------|
|         | Take Home      | Impulse     | Food Service   |
| Família | Potes 2 litros | Cornetto    | Potes 5 litros |
|         | Potes 1 litro  | Magnum      |                |
|         | Carte Dor      | Fruttare    |                |
|         | Ades           | Kids        |                |
|         |                | Clássicos   |                |
|         |                | Ades        |                |

A participação nas vendas totais de cada categoria segue a distribuição a seguir:

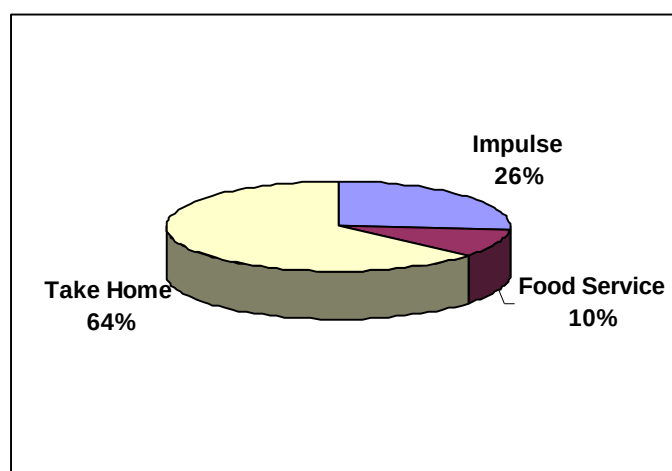


Figura 2.6 - Distribuição das vendas de 2004 por categoria

Fonte: Empresa

## 2.4 Áreas da empresa

A estrutura de áreas da Kibon está dividida em diversas diretorias e cada uma delas possui algumas gerências. Na figura 2.7, está destacada somente a área de *Supply Chain* e as respectivas partes que a compõem.

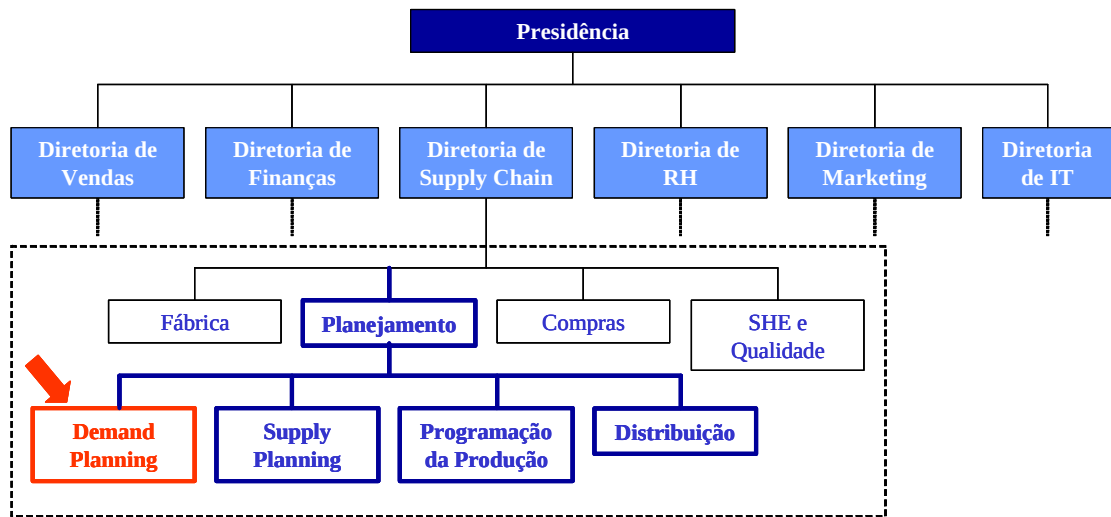


Figura 2.7 - Organograma da empresa e, em destaque, a área de Planejamento

Elaborado pela autora



## **O Processo de Planejamento**

A thick, horizontal orange brushstroke underline that starts with a light gradient on the left and becomes a solid orange on the right, ending in a small upward-pointing tail.

### 3. O PROCESSO DE PLANEJAMENTO

#### 3.1 Visão Geral

A área de Planejamento, como o próprio nome diz, planeja toda a demanda, necessidade de materiais, produção e distribuição dos produtos até os centros de distribuição.

O processo tem início com o planejamento da demanda, ou *Demand Planning*, no qual se estima o volume de vendas desde um horizonte maior (ano) até um horizonte médio (mês), num processo que conta também com a participação das áreas de Finanças, Trade Marketing e Marketing e que será explicado com maiores detalhes mais adiante.

É com base nesses números que se define o Plano Mestre de Produção (MPS), no qual é detalhado o volume mensal a ser produzido em cada uma das duas fábricas no nível de SKU (*Stock Keeping Unit*), ou seja, de cada um dos produtos da empresa.

Além da estimativa operacional definida inicialmente, neste momento são considerados diversos outros fatores, que impõem uma série de restrições com relação ao que poderá ou não ser produzido pelas fábricas. São eles:

- **Capacidade de produção:** engloba além da capacidade da própria máquina, a capacidade de mão-de-obra disponível.
- **Lote mínimo:** limite mínimo imposto para a produção de cada um dos produtos, visando reduzir o custo de produção.
- **Estoque disponível**
- **Prioridades:** a direção da empresa define estrategicamente uma prioridade entre os diversos produtos, uma vez que alguns deles compartilham suas produções em uma mesma máquina.

De posse do MPS, é preciso fazer o planejamento dos materiais necessários, tanto matéria-prima como material de embalagem, para atender a produção determinada. A equipe de *Supply Planning* verifica a quantidade em estoque dos materiais e a quantidade necessária e faz a interface com a equipe de compras.

O Plano Mestre de Produção precisa ser “quebrado” em um horizonte curto, nas semanas e nos dias de cada mês. O produto desta etapa é o chamado Plano de Produção (*Production Schedule*).

Na Kibon, somente a distribuição dos produtos das fábricas para os Centros de Distribuição (distribuição primária) e a transferência dos mesmos entre as fábricas compete à área de Planejamento. A chamada distribuição secundária (dos CDs para os pontos-de-venda) é de responsabilidade da área de *Customer Service*.

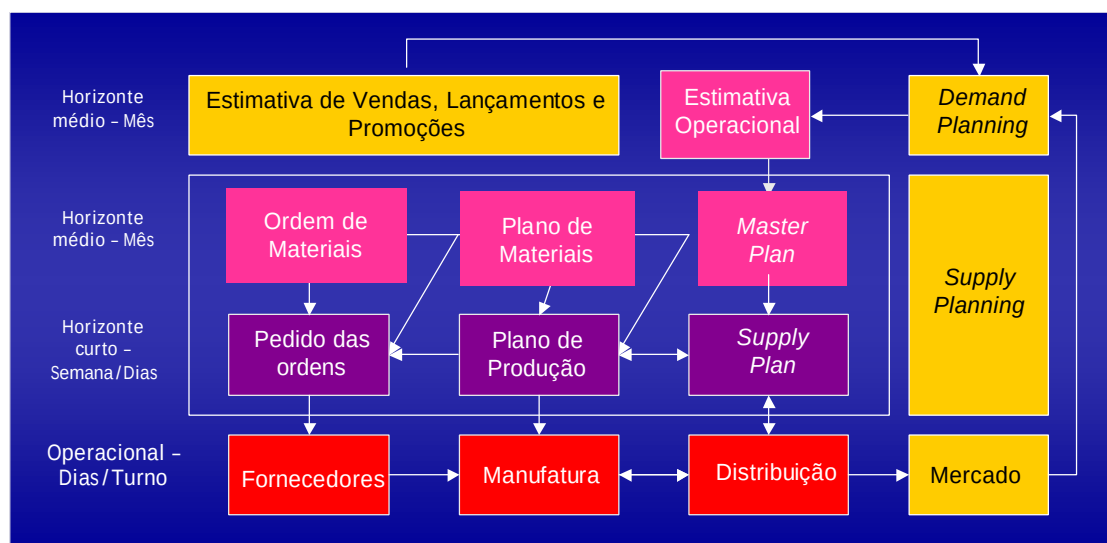


Figura 3.1 - Fluxograma do processo de Planejamento

Fonte: Empresa

## 3.2 Planejamento da Demanda

### 3.2.1 Componentes da Previsão

A previsão de demanda da empresa pode ser dividida da seguinte maneira:

- *Baseline Forecast*: componente da demanda baseada no histórico de vendas da companhia. É a base da previsão de demanda.
- *Event Forecast*: corresponde ao volume esperado devido aos eventos previstos, como promoções e propaganda, pelas áreas de marketing, trade marketing e vendas.
- *New Product Forecast*: corresponde a previsão do volume esperado para novos produtos a serem lançados. Avalia-se também a possível ocorrência de canibalização, ou seja, a diminuição de vendas de outro produto devido à entrada de um novo produto no mercado.

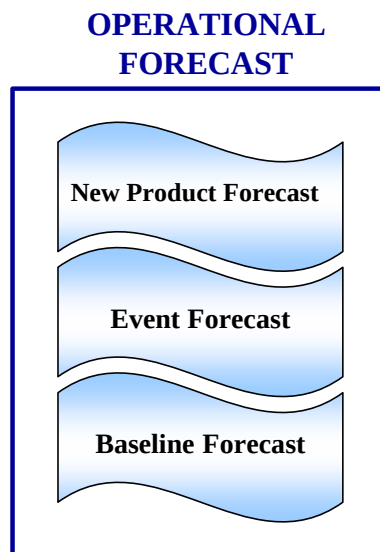


Figura 3.2 - Componentes da Previsão de demanda

Elaborado pela autora

### 3.2.2 Processo de Planejamento da Demanda

O processo de planejamento da demanda inicia-se com a definição do Plano Anual (AP) de trabalho pelos membros da diretoria de Finanças, Vendas, Marketing e *Supply Chain*. Neste plano, os volumes mensais de cada categoria (*Impulse*, *Take Home* e *Food Service*) são definidos levando-se em consideração o histórico de vendas reais de anos anteriores e a estratégia de negócio da empresa, sem o uso de qualquer modelo de previsão de demanda.

Dessa maneira, trabalha-se em um horizonte de médio prazo e num baixo nível de detalhamento, uma vez que este plano define somente os volumes de cada uma das duas regiões definidas pela empresa (Sul-Sudeste e Norte-Nordeste), área (Direta e Indireta), quebrados somente nas categorias.

Ao final de cada mês, é realizada então a chamada Reunião de Estimativa, na qual o volume do mês definido no AP é quebrado nos respectivos SKUs de cada categoria.

São realizadas, em geral, três reuniões, cada qual para uma categoria, lideradas pelo time de *Demand Planning* e com a participação de marketing e trade marketing. Cada uma dessas áreas é responsável por apresentar as informações relevantes para a definição dos volumes de cada um dos produtos como, por exemplo, as ações promocionais que estão sendo realizadas, ações de divulgação e o próprio histórico de vendas.

No entanto, o volume da categoria decidido no Plano Anual não pode ser modificado; sendo assim, trabalha-se em cima do chamado SOA – *Share of Assortment*, que nada mais é do que o percentual de participação de cada SKU no volume total de sua respectiva categoria, definido também com base no histórico de vendas, sem o uso formal de um método de previsão. Por exemplo: se está sendo feita uma promoção específica para o produto Picolé Fruttare Limão, uma maior parte do volume definido para a categoria *Impulse* é destinada a este produto através do aumento do seu SOA, em detrimento de algum outro SKU que, conseqüentemente, deverá ter seu SOA diminuído.

Apesar de não possuir um modelo bem definido de previsão de vendas, pode-se dizer que a empresa utiliza um método qualitativo, baseado no julgamento, na intuição, no conhecimento e, sobretudo, na experiência de seus diretores e gerentes. Outras características associadas a esse método citadas por BALLOU (2001), como dados históricos pouco relevantes, natureza

não-científica dos métodos, informações que afetam a previsão serem não-quantitativas e subjetivas, também são facilmente identificáveis na empresa.

O produto destas reuniões é o chamado Plano Operacional (OP), uma planilha com maior detalhamento do que o Plano Anual (pois entra no nível de SKU), com os volumes de cada produto revistos até o final do ano.



*Figura 3.3 - Processo de Planejamento da Demanda*

*Elaborado pela autora*

## **O Problema e Objetivo**

A thick, horizontal orange bar with a slight gradient, positioned below the title.

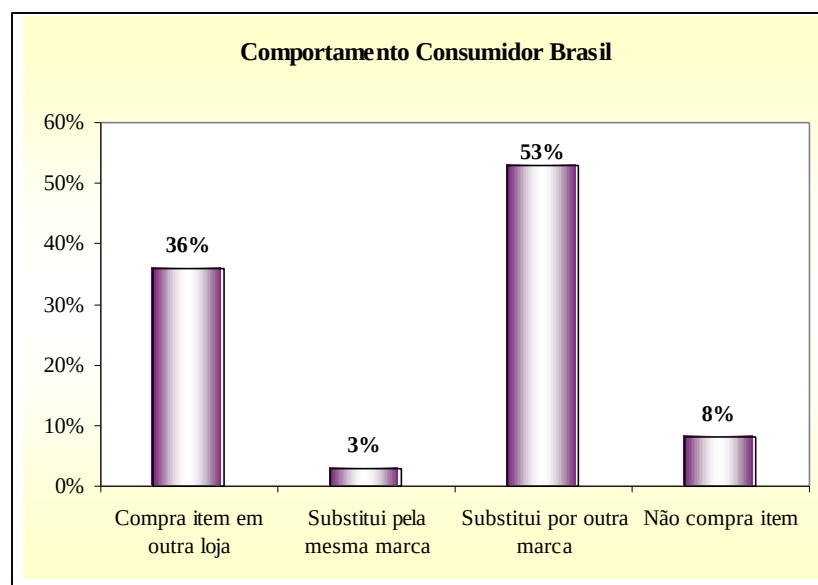
## 4. O PROBLEMA E OBJETIVO

### 4.1 Visão Geral

O mercado de bens de consumo vem passando por diversas mudanças nas últimas décadas, causadas pela globalização e o acirramento da concorrência entre as empresas. A grande variedade de produtos oferecidos aos consumidores faz com que os mesmos se tornem cada vez mais exigentes, levando diversos aspectos em consideração no momento da decisão pela compra de uma ou outra marca: preço, qualidade, disponibilidade e a própria força da marca, entre outros.

De acordo com pesquisa realizada pela AC Nielsen (2004), a falta do produto nos pontos de venda gera as seguintes reações por parte do consumidor:

- **53% substituem por outra marca**
- 36% compram o item em outra loja
- 8% não compram o item
- 3% substituem pela mesma marca



*Figura 4.1 - Comportamento do Consumidor Brasileiro*

*Fonte: AC Nielsen 2004*



Neste contexto, algumas áreas e operações das empresas merecem cada vez mais atenção por parte das mesmas, como é o caso da cadeia de suprimentos.

Percebe-se um crescimento dessa área nas empresas, e não somente daquelas produtoras de bens de consumo, uma vez que um planejamento de demanda, materiais, produção e distribuição bem executados podem proporcionar às empresas uma redução dos seus custos operacionais, além de, principalmente, aumentar ou garantir a disponibilidade de seus produtos nos diversos pontos de venda e conseqüentemente, conforme demonstrado através da pesquisa Nielsen da figura 4.1, as suas vendas.

HAX (1984) apresenta a definição dos seguintes elementos como sendo os mais impactantes para atingir esses objetivos, ou seja, entregar ao cliente o produto correto nas quantidades exatas, no local correto, com qualidade e a um custo competitivo:

- Múltiplas plantas
- Múltiplos centros de distribuição
- Múltiplos produtos
- Transporte e distribuição
- Comunicação e tecnologia de informação
- Pessoas e seus conhecimentos e habilidades

A aquisição e utilização dos elementos acima dependem de uma série de restrições, tais como de produção, de equipamentos, de mão-de-obra, além de estarem sujeitas às **incertezas da demanda**.

A tarefa de definir ou prever qual será efetivamente o volume de vendas ao longo do tempo não é fácil de ser realizada. Acredita-se que sejam diversos os fatores que exercem influência sobre a venda dos produtos, sejam eles fatores históricos, econômicos ou relacionados com a estratégia da empresa, como no caso de promoções e propaganda.

Estes últimos podem alterar significativamente os padrões de demanda. Dessa maneira, uma empresa que saiba mensurar o real impacto de suas ações estratégicas, sobretudo as citadas anteriormente, terá uma vantagem competitiva frente às suas concorrentes, uma vez que terá uma previsão de demanda mais precisa e poderá usufruir os benefícios trazidos por este fato a sua cadeia de suprimentos, os quais já foram apresentados anteriormente.

## 4.2 Definição do problema

A empresa enfrenta atualmente duas dificuldades que comprometem as suas operações. São elas:

- Falta de acuracidade das suas previsões de demanda, no que diz respeito ao *baseline forecast*
- Falta de conhecimento do impacto quantitativo exercido por diversos eventos, tanto internos como externos, sobre a venda de seus produtos (*event forecast*)

Para melhorar o *Baseline Forecast*, KIYUZATO (2001) aplicou três diferentes tipos de modelos de previsão (Suavização Exponencial, Série de Fourier e Série de Fourier “Suavizada”) e a combinação entre eles para alguns SKUs, classes e linhas de produtos diferentes, obtendo uma redução média de 25% dos desvios absolutos percentuais de previsão.

No entanto, os modelo aplicados, todos eles baseados em séries temporais, não contemplam o estudo dos fatores que exerceriam impacto nas vendas da empresa.

Conforme citado anteriormente, as áreas de vendas, marketing e trade marketing exercem um papel ativo na decisão dos valores do Plano Operacional, levando às reuniões de estimativa as informações a respeito das promoções, propagandas e outras ações que podem impactar nas vendas. No entanto, as decisões sobre o aumento ou diminuição do volume devido a essas ações são tomadas baseadas somente no conhecimento e na experiência dos profissionais envolvidos.

Dessa maneira, apesar do estudo de KIYUZATO (2001) ter melhorado o chamado *Baseline Forecast*, a falta de conhecimento a respeito dos eventos e da maneira de se quantificar o *Event Forecast* continuou sendo um problema para a previsão de demanda da Kibon.

### 4.3 Objetivos do trabalho

Tendo em vista o problema enfrentado pela empresa e citado na seção 4.2, o presente trabalho tem como objetivo levantar os principais fatores, internos e externos, que exerceriam influência sobre as vendas da Kibon, avaliar o grau do impacto existente e estudar a possível aplicação de um modelo de previsão de demanda.

Dessa maneira, espera-se possibilitar à empresa um maior conhecimento sobre os impactos exercidos pelas suas ações nas vendas de seus produtos e também por alguns fatores externos, e conseqüentemente disponibilizar uma ferramenta que auxilie na própria decisão pela realização ou não de uma ação, podendo por fim melhorar o *Event Forecast*.

Vale ressaltar neste momento que a autora realizou estágio na empresa durante o ano de 2004, na área de Planejamento, atuando em projetos relacionados à previsão de demanda, planejamento da produção e ao processo de abastecimento.

Durante o período de estágio, iniciou um estudo para o desenvolvimento de um modelo de previsão de demanda explicativo.

Para a realização deste estudo, contou com a colaboração dos funcionários da empresa, em especial da área de planejamento, obtendo acesso a todos os dados necessários, como os de volume de vendas, histórico de temperaturas, entre outros.

## **Revisão Bibliográfica**



## 5. REVISÃO BIBLIOGRÁFICA

### 5.1 Modelos de previsão

Os modelos de previsão de demanda podem ser classificados basicamente nas seguintes categorias, de acordo com MAKRIDAKIS (1998): qualitativos ou quantitativos. Essas classificações são detalhadas a seguir.

#### *Métodos qualitativos*

SANTORO (2003) refere-se a esse método como de *Predição* e o descreve como sendo um modelo aplicado em casos onde a demanda futura não possui relação com os padrões do passado e, portanto, não pode ser determinada com base nos mesmos. Dessa maneira, as previsões são geradas com base nos conhecimentos de algumas pessoas chaves, levando ou não em consideração a estratégia da empresa em questão.

#### *Métodos quantitativos*

MAKRIDAKIS (1998) aponta três condições que seriam as premissas para aplicação deste tipo de modelo. São elas:

- Informação sobre o passado está disponível
- Essa informação pode ser quantificada na forma de dados numéricos
- Algum padrão do passado continuará a existir no futuro

Os métodos quantitativos de previsão de demanda podem ainda ser classificados como de *Projeção (Temporais)* ou de *Explicação (Causais)*, dependendo se não há preocupação em se conhecer as causas que impactam a demanda (Projeção) ou se uma lei de dependência entre essas causas e a demanda futura é sabida e utilizada na previsão, numa relação de causa-e-efeito (Explicação).

A tabela 5.1 apresenta um resumo das características de cada um dos dois métodos acima, além dos principais modelos de cada um deles, que serão explicados com maiores detalhes posteriormente.

Tabela 5.1 - Métodos de previsão Quantitativos

|                    | Projeção / Temporais                                   |                        | Explicação / Causais                                                                     |                    |
|--------------------|--------------------------------------------------------|------------------------|------------------------------------------------------------------------------------------|--------------------|
| Características    | Não há preocupação em se conhecer as causas da demanda |                        | Demanda futura determinada levando-se em consideração as variáveis que impactam na mesma |                    |
| Principais modelos | Decomposição de séries temporais                       | Média Móvel            | Regressão Linear                                                                         | Regressão Simples  |
|                    |                                                        | Suavização Exponencial |                                                                                          | Regressão Múltipla |
|                    | Box-Jenkins                                            |                        |                                                                                          |                    |

De acordo com HAX (1984) é importante observar que mesmo nos casos de modelos de previsão baseados na matemática, a opinião dos especialistas e dos profissionais experientes responsáveis pela tomada de decisão não deve ser descartada; os resultados apresentados pelo modelo devem ser analisados e julgados pelos profissionais para se chegar aos valores finais de estimativa de demanda. WHEELWRIGHT (1977) apresenta visão semelhante, ressaltando o fato de que os métodos de previsão de demanda devem ser uma ferramenta de auxílio da tomada de decisão e não a decisão propriamente dita.

Alguns aspectos devem ser levados em consideração ao se fazer uma previsão de demanda. WHEELWRIGHT (1997), apresenta o que seriam os três elementos básicos da previsão:

1. Tempo: a análise da demanda está relacionada, direta ou indiretamente, ao tempo.
2. Incerteza: a incerteza está presente em qualquer situação em que haja uma previsão.
3. Confiabilidade dos dados: os métodos de previsão de demanda estão baseados na existência de dados históricos confiáveis.

## **5.2 Etapas de aplicação de um modelo de previsão**

MAKRIDAKIS (1998) define como sendo cinco as etapas do processo de aplicação de um modelo de previsão quantitativo.

### **Etapa 1: Definição do problema**

Essa fase consiste no levantamento de informação a respeito da necessidade do uso de um modelo de previsão, a quem e a qual processo ele seria destinado. Os dados relacionados ao produto ou serviço que será previsto também devem ser levantados, pois impactam na decisão de qual modelo deve ser aplicado.

### **Etapa 2: Levantamento de informações**

Levantamento dos dados históricos propriamente ditos. Esses dados serão utilizados no modelo escolhido.

### **Etapa 3: Análise preliminar dos dados**

Os dados levantados na etapa anterior devem ser analisados, ainda de uma maneira visual, através de gráficos, para que a tendência, sazonalidade e outras características da demanda possam ser identificadas.

É nesta etapa que ocorre uma primeira identificação dos possíveis modelos a serem utilizados.

### **Etapa 4: Escolha e ajuste dos modelos**

Nesta etapa, os modelos a serem testados são escolhidos com base em todas as informações e análises levantadas anteriormente. Os modelos são então testados para que um deles possa ser escolhido ao final, levando em consideração os resultados por eles apresentados.

### **Etapa 5: Uso e avaliação do modelo**

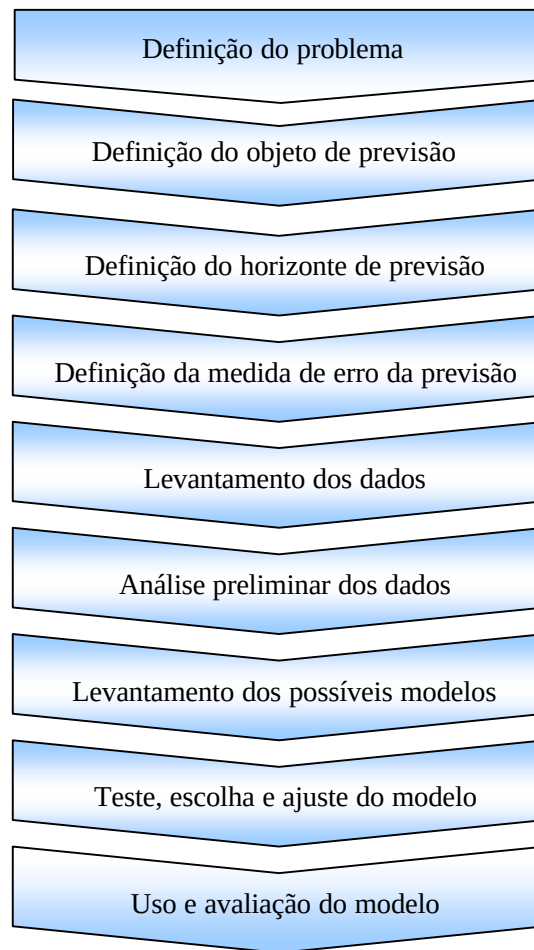
Após a escolha do modelo, ele deve ser testado e validado com as pessoas que passarão a utilizá-lo ou por aquelas que irão utilizar seus resultados. Vale ressaltar que um modelo só poderá ser completamente aprovado quando forem efetivamente realizadas previsões de períodos futuros, e não somente dos dados históricos levantados.

DIEBOLD (1997) apresenta uma abordagem semelhante, também dividindo o processo inicial de aplicação de um modelo de previsão em etapas também. As etapas definidas pelo autor são, de uma maneira geral, as seguintes:

1. **Ambiente de decisão:** definição de quais serão as decisões da companhia que serão guiadas pela previsão obtida e como será medida a qualidade da mesma.
2. **Objeto da previsão:** definição de qual a série que se deseja prever e levantamento dos dados a respeito da mesma (que tipo de série que se trata, qual o período de dados que está disponível, entre outros).
3. **Declaração da previsão:** como os valores de previsão serão entregues. Neste momento se definiria, por exemplo, se há necessidade de um intervalo de possíveis valores de previsão ou se bastaria somente um valor.
4. **Horizonte de previsão:** definição de qual o horizonte a ser previsto, por exemplo, um mês, dois meses, um ano ou outro período.
5. **Levantamento das informações:** levantamento dos dados que serão utilizados no modelo de previsão.
6. **Definição do modelo de previsão:** escolha do modelo a ser utilizado levando em consideração a complexidade do mesmo, a aplicação e a adequação aos dados disponíveis.

Levando em consideração as etapas definidas por ambos os autores, pode-se resumir o processo de escolha e aplicação de um modelo de previsão da maneira apresentada na figura 5.1.





*Figura 5.1 - Etapas para escolha e aplicação dos modelos de previsão*

*Elaborado pela autora*

### 5.3 Decomposição de séries temporais

A decomposição de séries temporais parte do princípio de que o padrão de demanda pode ser quebrado em quatro principais componentes: tendência, ciclicidade, sazonalidade e componente aleatória. No entanto, segundo MAKRIDAKIS (1998), em muitos modelos de decomposição, as componentes tendência e ciclicidade são consideradas uma só componente (a de tendência-ciclo) por considerar a distinção entre elas um tanto quanto artificial.

A principal vantagem de se decompor os padrões passados da demanda consiste no fato de que, dessa maneira, pode-se melhor entender o comportamento da demanda: as componentes de tendência e ciclo estariam relacionadas ao comportamento a longo prazo da série, enquanto a sazonalidade explica as variações encontradas durante o período completo em estudo.

De maneira geral, existem duas abordagens diferentes para a decomposição de séries temporais: a decomposição aditiva e a multiplicativa.

Na primeira delas, a **aditiva**, o valor a ser previsto é determinado a partir da soma das componentes apresentadas anteriormente. De acordo com MAKRIDAKIS (1998), esse tipo de decomposição é aplicado em casos em que a magnitude da componente sazonal não varia de acordo com o nível da série e pode ser representada pela equação (1):

$$Y_t = S_t + T_t + E_t \quad (1)$$

onde,  $Y_t$  é o valor observado da série no período  $t$ ,

$S_t$  é a componente sazonal no período  $t$ ,

$T_t$  é a componente tendência no período  $t$  e

$E_t$  é a componente aleatória no período  $t$ .

Já a decomposição **multiplicativa** é utilizada nos casos em que a componente sazonal sofre variações de acordo com o aumento ou a diminuição no nível das séries e é representada pela equação (2).

$$Y_t = S_t \times T_t \times E_t \quad (2)$$

As abordagens apresentadas a seguir são baseadas então neste princípio da decomposição de séries temporais.

## 5.4 Médias

### 5.4.1 Média Simples

É o método mais simples de previsão e consiste em utilizar a média de todos os dados históricos como previsão. Dessa maneira,

$$F_{t+1} = \frac{1}{t} \sum_{i=1}^t Y_i \quad (3)$$

onde,  $F_{t+1}$  é a previsão do período  $t+1$ .

De acordo com MAKRIDAKIS (1998), o método de média simples é apropriado para séries onde não há tendência nem sazonalidade, uma vez que o cálculo da previsão é baseado em muitos valores de dados históricos.

### 5.4.2 Média Móvel

O termo “móvel” deve-se ao fato de que, a cada nova média calculada, o dado mais antigo é descartado e um mais novo é adicionado ao cálculo da média, mantendo em seu cálculo o mesmo número de observações.

$$F_{t+1} = \frac{1}{k} \sum_{i=t-k+1}^t Y_i \quad (4)$$

Na fórmula acima,  $k$  é a ordem da média, ou seja, o número de dados utilizado no cálculo da previsão.

Segundo MAKRIDAKIS (1998), esse método também não é adequado para casos em que se percebe a presença de tendência ou sazonalidade, apesar de lidar melhor com essas componentes do que o método de média simples.

## 5.5 Regressão Linear

Os métodos de regressão linear estão dentro do grupo dos chamados métodos explicativos, uma vez que relacionam a variável a ser prevista, a variável dependente, com outras variáveis, as independentes.

Nos casos em que se deseja prever uma determinada variável  $Y$ , relacionando-a a uma variável  $X$  e assim desenvolvendo um modelo que relacione  $Y$  e  $X$ , temos o que se chama de Regressão Linear Simples.

Em outros casos, o comportamento de uma variável  $Y$  pode ser explicada por mais de uma variável,  $X_1, X_2, \dots, X_k$ . Neste caso, temos o que se chama de Regressão Linear Múltipla.

KACHIGAN (1986) aponta como sendo os principais objetivos da Regressão Linear:

1. Determinar se existe ou não uma relação entre duas variáveis.
2. Descrever essa relação, quando existir, na forma de uma equação matemática.
3. Possibilitar o conhecimento do grau de acuracidade obtido a partir da equação de regressão.
4. Determinar a importância relativa de cada uma das variáveis de previsão na previsão da variável dependente.

Dessa maneira, a utilização do modelo de regressão linear está baseado no fato de uma ou mais variáveis exercerem influência na variável que se deseja prever. Para se entender então qual o grau de relacionamento entre elas, ou mesmo se existe esse relacionamento, é importante analisar o coeficiente de correlação existente.

Sendo assim, a primeira parte dessa seção será dedicada à explicação do coeficiente de correlação.

### 5.5.1 Coeficiente de correlação

De acordo com KACHIGAN (1986), apesar de gráficos, relacionado o comportamento de duas variáveis, darem uma primeira idéia do grau de influência de uma sobre a outra, é desejável poder quantificar esse grau de associação.

Uma das maneiras de se fazer isso é através do Coeficiente de Correlação, designado pela letra  $r$ , quando se tratar de uma amostra, e pela letra grega  $\rho$  quando se referir à população, podendo variar de -1 a +1.

De acordo com KACHIGAN (1986), um coeficiente igual a +1 indica uma correlação positiva perfeita, enquanto um coeficiente igual a -1 indica uma correlação negativa perfeita. No entanto, raramente um coeficiente de correlação apresentaria um desses valores extremos.

Ainda de acordo com o mesmo autor, interpretando esse coeficiente de uma maneira mais profunda, pode-se dizer que um valor de  $r = 0,8$  significa que 0,64 (0,8 ao quadrado), ou seja, 64% da variação de uma das variáveis está relacionada com a variação da outra.

De uma maneira geral, MAKRIDAKIS (1998) resume como sendo duas as evidências que esse coeficiente indicaria:

1. O sinal, positivo ou negativo, indica a direção da relação entre as duas variáveis. Uma correlação positiva indica que a variação de uma delas será acompanhada pela outra no mesmo sentido; isto é, se uma aumenta, a outra aumenta também; se uma diminui, a outra variável responde também com uma diminuição do seu valor. Por outro lado, uma correlação igual a zero indica que não há uma relação entre as duas variáveis.
2. Quanto maior o valor, em módulo, do coeficiente, maior é a correlação entre as duas variáveis.

A seguir, temos as equações que permitem o cálculo do coeficiente de correlação:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (5)$$

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i \quad (6)$$

$$S_x^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (7)$$

$$S_y^2 = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2 \quad (8)$$

$$r_{xy} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{(n-1)S_x S_y} \quad (9)$$

onde  $\bar{X}$  é a média dos valores de  $X$ ,

$\bar{Y}$  é a média dos valores de  $Y$ ,

$S_x$  é o desvio padrão da variável  $X$  e

$S_y$  é o desvio padrão da variável  $Y$

Por fim, vale ressaltar que o uso do coeficiente de correlação deve ser acompanhado de alguns cuidados, de acordo com KACHIGAN (1986):

1. O coeficiente de correlação é apropriado apenas para os casos em que as variáveis possuem uma relação linear. Se elas possuírem uma relação, mas uma não-linear, mesmo assim o coeficiente será zero.
2. Os dados utilizados para seu cálculo devem ser aleatórios.
3. As duas variáveis devem ter uma distribuição normal.

No entanto, MAKRIDAKIS (1998) ainda aponta outros dois cuidados a serem tomados:

4. O tamanho da amostra não pode ser pequeno. Em geral, uma amostra em que  $n = 30$ , ou acima disso, é desejável.
5. Na amostra, não deve haver a presença de um valor extremo, ou seja, a presença de um *outlier*, pois isso pode afetar o valor do coeficiente de correlação.

### 5.5.2 Regressão Linear Simples

Como explicado anteriormente, a Regressão Linear Simples relaciona uma determinada variável dependente  $Y$  com uma única variável  $X$ , explicada numa relação:

$$Y = a + bX + e \quad (10)$$

onde  $a$  é uma constante,

$b$  se refere à inclinação da reta e

$e$  é o erro.

É preciso então saber como se determinar os valores de  $a$  e  $b$  de maneira que a reta encontrada se ajuste da melhor maneira ao conjunto de dados apresentados.

Para isso, pode-se utilizar o método dos mínimos quadrados, que consiste em achar os valores de  $a$  e  $b$  de maneira que a soma do erro, ou seja, a diferença entre o valor previsto e o valor real ao quadrado seja minimizada.

Os valores de  $b$  e  $a$  que minimizam então a soma do erro ao quadrado são:

$$b = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} \quad (11)$$

$$\text{e } a = \bar{Y} - b\bar{X} \quad (12)$$

Uma vez encontrada a equação da regressão e o relacionamento entre as duas variáveis, os valores de  $Y$  podem ser comparados aos valores obtidos a partir da reta de regressão, chamados de  $\hat{Y}$ .

Dessa maneira, a correlação entre  $Y$  e  $\hat{Y}$  é dado por:

$$R^2 = r_{Y\hat{Y}}^2 = \frac{\sum (Y_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2} \quad (13)$$

Para determinar se a equação de regressão obtida é adequada ou não para representar o relacionamento das duas variáveis, pode-se traçar um gráfico de pontos, relacionado erros  $e_i$ , ou chamados **resíduos**, em relação à variável independente  $X_i$ .

De acordo com MAKRIDAKIS (1998), para se considerar que a equação de regressão é adequada para os dados em questão, não deve haver relação entre o resíduo e a variável  $X$ ; isso ocorre quando o gráfico de pontos não apresenta nenhum tipo de tendência, como curvas ou crescimentos. Os pontos devem estar distribuídos aleatoriamente nos dois lados do eixo horizontal, ou seja, tanto abaixo de zero quanto acima.

No entanto, também é preciso verificar antes a presença de *outliers*. *Outliers* são pontos que apresentam grande resíduo, distantes dos demais pontos, e representam os dados que por algum motivo não se ajustam à reta determinada pela equação de regressão.

Em muitos casos, somente a análise visual do gráfico de resíduos não permite que se afirme com certeza a existência da relação entre duas variáveis. Para verificar tal relação, é necessário então a aplicação de testes estatísticos, como os a seguir, apresentados por MAKRIDAKIS (1998).

### Teste- $F$

O teste- $F$  é um dos testes estatísticos que permite avaliar se há uma relação significativa entre as variáveis analisadas, e é dado por:

$$F = \frac{\sum (\hat{Y}_i - \bar{Y})^2 / (m - 1)}{\sum (Y_i - \hat{Y}_i)^2 / (n - m)} \quad (14)$$

onde  $m-1$  é o chamado **grau de liberdade** e  $m$  é o número de coeficientes na equação de regressão, que no caso de uma regressão simples é igual a 2.

Uma outra maneira de se escrever  $F$  é a apresentada na equação (15), que relaciona  $F$  com o coeficiente de correlação entre o valor real e o valor previsto.

$$F = \frac{R^2 / (m - 1)}{(1 - R^2) / (n - m)} \quad (15)$$

Caso haja mesmo uma relação entre as duas variáveis, espera-se que a parte da variação que não seja explicada pela regressão, ou seja os erros, seja pequena e, portanto, o valor de  $F$  seja grande.

A partir do valor de  $F$ , obtém-se o valor  $P$ , que seria a probabilidade de se obter um valor de  $F$  como o encontrado se o valor de  $b$ , ou seja, a tendência, for zero. MAKRIDAKIS (1998) aponta que valores de  $P$  abaixo de 0,05 indicariam a existência de uma regressão relevante.

Na aplicação prática da regressão linear para a previsão de demanda, espera-se obter um certo valor de  $Y$ , dado um valor  $X_0$ ; dessa maneira, de acordo com a equação (16) teremos:

$$\hat{Y}_0 = a + bX_0 \quad (16)$$



Sabe-se que o valor previsto de  $Y$  através da equação (16) será diferente do valor real. Sendo assim, é necessário definir o erro padrão da previsão, dado pela equação (17).

$$s.e.(Y_o) = s_e \sqrt{1 + \frac{1}{n} + \frac{(X_o - \bar{X})^2}{\sum (X_i - \bar{X})^2}} \quad (17)$$

Assim, a previsão obtida através da regressão será dada pelo valor de  $Y$  encontrado, somado ao erro obtido pela equação acima multiplicado por  $t^*$ .

$$Y_o \pm t^* s.e.(Y_o) \quad (18)$$

O valor de  $t^*$  é obtido através de uma tabela, tendo em vista o intervalo de confiança definido e os graus de liberdade, dado por  $n-m$ , como no cálculo de  $F$  apresentado anteriormente. De uma maneira geral, quanto maior o intervalo de confiança estipulado maior será o valor de  $t^*$  e, portanto, o intervalo de variação de  $Y$ .

### 5.5.3 Regressão Linear Múltipla

Ao contrário da regressão simples, na regressão múltipla o comportamento de uma variável é explicado por duas ou mais variáveis independentes.

$$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + \dots + b_k X_k + e \quad (19)$$

Os coeficientes da equação podem ser encontrados através do método dos mínimos quadrados, ao se buscar aqueles coeficientes que minimizam a soma dos erros ao quadrado, como no caso da regressão linear simples.

### **Coeficiente de correlação múltipla e teste $F$**

Assim como na Regressão Linear Simples, o grau de relacionamento entre as variáveis de explicação e a variável dependente pode ser medido através do coeficiente de correlação, do teste  $F$  e do valor  $P$ .

O chamado coeficiente de correlação múltipla é dado por:

$$R = r_{yy}^{\wedge} = \frac{\sum_{i=1}^n (Y_i - \bar{Y})}{(Y_i - \bar{Y})} \quad (20)$$

onde  $\hat{Y}$  é o valor obtido através da equação de regressão.

Já o valor do teste  $F$  pode ser encontrado através da equação (21).

$$F = \frac{R^2 / k}{(1 - R^2) / (n - k - 1)} \quad (21)$$

onde  $k$  é o número de variáveis de explicação.

Importante ressaltar que todos os indicadores apresentados medem se há relação entre a variável  $Y$  e as variáveis independentes, não mostrando o grau de relacionamento entre  $Y$  e cada uma das variáveis separadamente.

### Teste $t$

O teste  $t$  mede o grau de significância de cada um dos coeficientes da equação de regressão, considerando a presença das demais variáveis de explicação. O valor  $t$  é dado por:

$$t = \frac{b_j}{se(b_j)} \quad (22)$$

onde  $b_j$  é o  $j$ -ésimo coeficiente e

$se(b_j)$  é o erro padrão de  $b_j$ .

A partir deste valor  $t$  pode-se obter o valor  $P$  para verificar o efeito de cada uma das variáveis.

A análise individual de cada valor  $t$  parte do princípio de que os coeficientes respectivos a cada uma das variáveis de explicação não são correlacionados entre si.

De acordo com MAKRIDAKIS (1998), a aplicação destes testes apresentados anteriormente está condicionada a:

- Escolha correta do modelo (por exemplo, linear ou não linear)
- Resíduos independentes, ou seja, não correlacionados entre si

- Homocedasticidade: resíduos com variância constante
- Distribuição normal dos resíduos

Para o caso da aplicação da regressão múltipla em séries temporais, MAKRIDAKIS (1998) sugere outros cuidados a serem tomados. São eles:

- Checar a independência entre os resíduos: isso pode ser feito através da análise da função de autocorrelação (FAC) dos resíduos.
- Se necessário, adicionar à equação de regressão variáveis que representem os efeitos relacionados com o tempo. Por exemplo, quando há a presença de sazonalidade na série, variáveis do tipo *dummy*, ou binárias, podem ser utilizadas, assumindo os valores 0 ou 1, dependendo do período em questão.
- A realização de uma previsão através da equação de regressão linear depende da existência prévia de previsão para os valores de cada uma das variáveis independentes.

Pode-se concluir, a partir deste último aspecto, que a escolha das variáveis que irão compor o modelo é de extrema importância e MAKRIDAKIS (1998) apresenta alguns métodos para a determinação destas variáveis.

A partir de uma lista levantada pelos especialistas envolvidos no processo, a escolha do melhor modelo, e conseqüentemente das variáveis que o compõe, poderia ser feita através daquele que apresentasse o maior valor de  $R^2$ . No entanto, o melhor modelo seria aquele que contivesse todas as variáveis pré-selecionadas, uma vez que a adição de uma variável à equação de regressão necessariamente aumenta o  $R^2$ .

Sendo assim, uma alternativa seria a análise do chamado  $R^2$  ajustado, dado por:

$$\bar{R}^2 = 1 - (1 - R^2) \frac{n-1}{n-k-1} \quad (23)$$

onde  $n$  é o número de observações e

$k$  é o número de variáveis de explicação do modelo.

O valor  $R^2$  ajustado deve ser calculado para todas as opções possíveis de equação e, portanto, este método só é viável quando o número de variáveis independentes consideradas for pequeno.

Quando isto não for possível, pode-se utilizar o método chamado de *stepwise*, que por sua vez é dividido em três tipos:

- *Stepwise forward*: calcula-se os resíduos da regressão contendo inicialmente somente aquela variável com maior correlação com a variável dependente  $Y$ . Depois, escolhe-se a variável com maior correlação com esses resíduos encontrados e assim sucessivamente, até que não haja mais qualquer variável com correlação significativa com os últimos resíduos calculados.
- *Stepwise backward*: neste método, inicia-se com a equação de regressão contendo todas as variáveis e depois deve-se ir retirando aquelas que apresentam os coeficientes com menor valor  $t$ .
- *Stepwise forward-with-a-backward-look*: primeiro, encontra-se a melhor variável de explicação, depois o melhor par de variáveis e por fim o melhor trio. A partir desse ponto, analisa-se se a última variável incluída no modelo deve ser removida ou não através de critérios definidos, como por exemplo, o valor  $R^2$  ajustado.

### Aplicação da Regressão Múltipla na Previsão

Para se medir a qualidade das previsões geradas pela equação de regressão múltipla, pode-se calcular o erro padrão, dado por:

$$\boxed{s.e.(\hat{Y}_0) = se \sqrt{1 + \mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}}} \quad (24)$$

onde  $\mathbf{c}$  é o vetor dos novos valores das variáveis de explicação e

$\mathbf{X}$  é a matriz de ordem  $n$  por  $k+1$ , na qual a primeira coluna é de valores 1 e as demais são vetores das variáveis de explicação.

## 5.6 ARIMA

Os modelos ARIMA (*Autoregressive Integrated Moving Average*), também conhecidos como método de Box-Jenkins, são normalmente utilizados para séries de tempo complexas, onde não há um padrão aparente de comportamento, conforme citado por WHEELWRIGHT (1977).

Assim como nos demais métodos de previsão, no ARIMA deve ser realizada uma análise preliminar dos dados e escolher o melhor modelo aplicável aos mesmos. No entanto, neste caso, são três os aspectos que devem ser analisados para que os parâmetros do modelo possam ser definidos.

A seguir, será detalhada cada uma dessas partes que compõem o modelo ARIMA.

### Autocorrelação e autocorrelação parcial

O coeficiente de autocorrelação pode ser calculado a partir da fórmula (9), apresentada na seção 5.5.1, e indica a existência ou não de uma relação entre o valor observado em **um certo instante  $t$  com os valores de instantes anteriores.**

$$r_k = \frac{\sum_{t=k+1}^n (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2} \quad (25)$$

onde  $k$  indica a ordem.

Um coeficiente de ordem 1 indica como valores sucessivos da série se relacionam; de ordem 2 indica como valores separados por dois períodos de tempo se relacionam, e assim sucessivamente.

O coeficiente de autocorrelação parcial é usado para se medir o grau de associação entre  $Y_t$  e  $Y_{t-k}$ , removendo os efeitos de outros períodos de tempos  $-1, 2, 3, \dots, k-1$ .

MAKRIDAKIS (1998) explica que se houver relação entre  $Y_t$  e  $Y_{t-1}$ , conseqüentemente haverá relação entre  $Y_{t-1}$  e  $Y_{t-2}$ , uma vez que esses dados também estão separados por somente um período de tempo. Dessa maneira, haverá relação também entre  $Y_t$  e  $Y_{t-2}$ .

**O coeficiente que mede a relação existente entre  $Y_t$  e  $Y_{t-k}$ , sem a influência dos valores  $Y$  de outros períodos é chamado de coeficiente de autocorrelação parcial.**

O coeficiente de autocorrelação parcial  $a_k$  pode ser encontrado a partir da chamada equação de autoregressão (AR), que difere da equação de regressão pelo fato das variáveis de explicação serem os próprios valores anteriores de  $Y$ .

$$\boxed{Y_t = b_0 + b_1 Y_{t-1} + b_2 Y_{t-2} + \dots + b_k Y_{t-k} + e_t} \quad (26)$$

O coeficiente  $a_k$  é o coeficiente  $b_k$  da equação de regressão múltipla (19).

De acordo com MAKRIDAKIS (1998), o coeficiente de autocorrelação para séries não-estacionárias, ou seja, séries com uma tendência de crescimento ou declínio, apresentam valores diferentes de zero durante diversas ordens, até que a participação e a influência do componente de erro aleatório passe a dominar a autocorrelação.

Sendo assim, deve-se remover este efeito apresentado por séries não-estacionárias e, para tanto, MAKRIDAKIS (1998) apresenta o método de *diferenciação*.

Na primeira diferenciação, ou diferenciação de ordem 1, o novo  $Y_t$  será dado pela diferença entre o  $Y_t$  real e o  $Y_{t-1}$ .

$$\boxed{Y'_t = Y_t - Y_{t-1}} \quad (27)$$

Esse procedimento deve ser aplicado sucessivamente, até que a série representada pelos novos valores  $Y'_t$  apresente comportamento estacionário.

### **Média Móvel**

De maneira análoga à equação (26), pode-se determinar uma equação de autoregressão tendo os erros de períodos passados como variáveis de explicação, num modelo chamado de média móvel (MA).

$$\boxed{Y_t = b_0 + b_1 e_{t-1} + b_2 e_{t-2} + \dots + b_k e_{t-k} + e_t} \quad (28)$$

MAKRIDAKIS (1998) atenta para o fato de que o termo média móvel deste caso não deve ser confundido com aquele apresentado na seção 5.4, que tratava das médias dos valores **observados**.

O modelo que associa as duas regressões expressas pelas equações (26) e (28), de autoregressão e de média móvel, na análise de séries não-estacionárias (com a presença de tendência ou sazonalidade), é o chamado modelo ARIMA (*Auto Regressive Integrated Moving Average*).

Três coeficientes definem este modelo. São eles:

- AR:**  $p$  = ordem da parte relativa à autoregressão  
**I:**  $d$  = grau da diferenciação envolvida  
**MA:**  $q$  = ordem da parte relativa à média móvel

De uma maneira geral, um modelos ARIMA do tipo  $(p,0,0)$ , ou seja, onde o valor de  $Y_t$  depende somente dos valores anteriores de  $Y$ , pode ser escrito da maneira apresentada na fórmula (29).

$$Y_t = c + f_1 Y_{t-1} + f_2 Y_{t-2} + \dots + f_p Y_{t-p} + e_t \quad (29)$$

onde  $c$  é uma constante,

$f_j$  é j-ésimo parâmetro de autoregressão e

$e_t$  é o erro no período  $t$ .

Já a equação (30) apresenta a fórmula para os modelos ARIMA  $(1,1,1)$ .

$$(1 - f_1 Y_{t-1})(1 - Y_{t-1})Y_t = c + (1 - q_1 e_{t-1})e_t \quad (30)$$

onde  $q_1$  é o primeiro parâmetro de média móvel.

Devido a grande variedade de modelos ARIMA existentes, MAKRIDAKIS (1998) sugere que as seguintes etapas sejam seguidas para se determinar o modelo mais adequado:

1. Plotar os dados e identificar a existência de observações incomuns. Caso necessário, ajustar os dados para estabilizar a variância.
2. Verificar se os dados são estacionários. Se a série se mantém constante ao longo do tempo em torno de um eixo horizontal médio, então a série é estacionária. Analisar também a função de autocorrelação (FAC) e a de autocorrelação parcial (FACP). Se a FAC e a FACP caírem para zero ou valores próximos a zero rapidamente, isso indica que os dados são estacionários.
3. Quando os dados não forem estacionários, torná-los através da diferenciação. Ela deve ser de ordem um quando não houver a presença de sazonalidade, e do tipo sazonal quando sua presença for constatada. Caso ainda assim os dados permanecerem não-estacionários, realizar nova diferenciação.
4. Depois que os dados estiverem estacionários, examinar a função de autocorrelação para verificar a existência de padrões:
  - A aplicação de modelos com a sazonalidade é sugerida quando a autocorrelação ou a autocorrelação parcial for significativamente diferente de zero nos pontos referentes ao período de sazonalidade.
  - Caso não haja valores significativos de autocorrelação após uma determinada ordem  $q$ , então um modelo de média móvel,  $MA(q)$ , deve ser adequado. Caso não haja uma autocorrelação parcial significativa para ordens maiores do que  $p$ , um modelo de auto-regressão,  $AR(p)$ , deve ser o mais apropriado.
  - Caso os dados não indiquem claramente se um modelo AR ou um MA é o mais adequado, uma mistura desses modelos deverá ser aplicada.



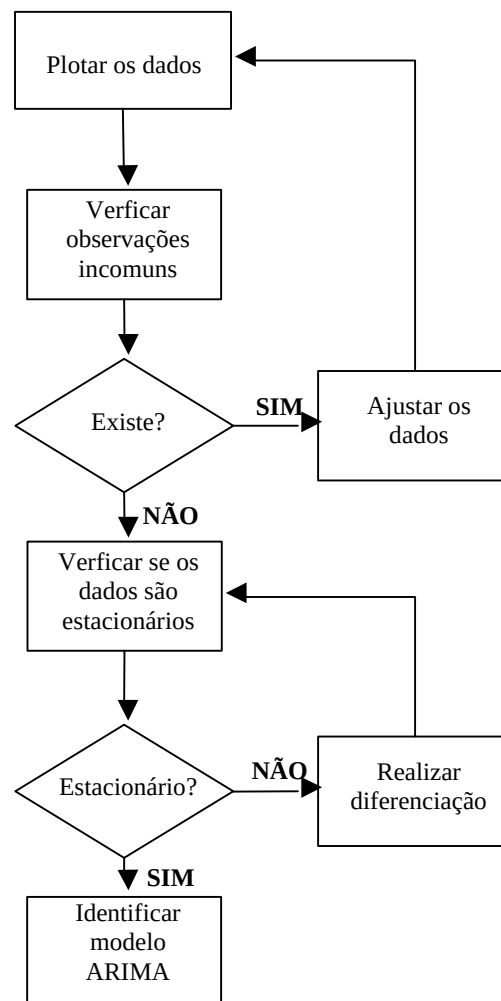


Figura 5.2 - Modelo para identificação do modelo ARIMA adequado

RAMOS (1999) resume os comportamentos das funções FAC e FACP que definiriam a escolha do modelo mais adequado, conforme a tabela 5.2.

Tabela 5.2 - FAC e FACP para identificação do modelo ARIMA

| Modelo    | FAC                   | FACP                  |
|-----------|-----------------------|-----------------------|
| $(p,0,0)$ | diminui gradualmente  | se anula após ordem p |
| $(0,0,q)$ | se anula após ordem q | diminui gradualmente  |
| $(p,0,q)$ | diminui gradualmente  | diminui gradualmente  |

Após a escolha e aplicação de um modelo, sua adequação pode ser verificada através da função de autocorrelação dos resíduos gerados; caso as diferenças entre os valores reais e aqueles gerados pela equação do modelo não estiverem autocorrelacionados, então o modelo escolhido pode ser considerado adequado, conforme apresentado por RAMOS (1999).

## **Definição e Levantamento de Dados**



## **6. DEFINIÇÃO E LEVANTAMENTO DOS DADOS**

### **6.1 Considerações iniciais**

Para o estudo em questão podem ser destacadas duas análises como sendo fundamentais:

1. Levantamento e avaliação do impacto de fatores que exerceriam influência sobre as vendas dos produtos através de métodos estatísticos
2. Estudo de aplicação de um modelo de previsão a partir dos fatores estudados.

Para tanto, é necessário definir claramente alguns pontos importantes, tendo como base as etapas para aplicação de um modelo de previsão, apresentadas na seção 5.2. São eles:

- Definição do problema (já apresentado na seção 4.2)
- Definição do objeto de estudo e previsão
- Definição do período estudado e horizonte de previsão

A seguir, cada um desses dois últimos pontos e as partes que o compõem serão definidos com maiores detalhes.

### **6.2 Objeto de estudo e previsão**

#### *6.2.1 Objeto de previsão*

O objeto, cuja previsão busca-se melhorar, é o volume de vendas da empresa.

A empresa possui um banco de dados com o histórico do volume de vendas desde 2002. Neste sistema, as informações podem ser extraídas das seguintes maneiras:

- faturado bruto ou líquido
- por ano, quadrimestre, mês ou diário
- por filial
- por família, grupo de produto ou produto

O volume de vendas a ser considerado neste trabalho será o faturamento líquido.

### 6.2.2 Fatores a serem estudados

Outra definição de extrema importância para a realização deste trabalho é com relação a quais fatores seriam aqueles que poderiam exercer algum impacto nas vendas da empresa.

É importante observar que os fatores levantados a seguir representam a opinião dos profissionais da empresa, baseados somente em suas experiências, e não se pode afirmar neste momento que estes representem realmente todos os fatores que exercem influência na demanda da Kibon.

É objetivo deste trabalho verificar a veracidade de tais suposições, analisando com base em métodos e modelos estatísticos a real existência de relação entre os fatores apontados pela empresa. São eles:

- **Temperatura:** sendo o sorvete um produto cujas vendas apresentam comportamento sazonal, com vendas significativamente maiores nos períodos mais quentes do ano, acredita-se que este seja o fator que mais exerceria influência nas vendas da Kibon.
- **Propaganda:** a veiculação de propaganda na televisão e outros meios de comunicação, como o rádio, a Internet, revistas e tablóides, é uma das ferramentas utilizadas pela empresa na tentativa de aumentar o seu volume de vendas.
- **Promoções:** espera-se que a realização de promoções nos pontos de venda gere um aumento significativo das vendas da Kibon. A distribuição de brindes na compra de um sorvete é um exemplo de promoções realizadas pela empresa.
- **Descontos:** os descontos são oferecidos pela área de vendas normalmente para os grandes clientes da empresa; podem ser descontos oferecidos diretamente sobre o valor de venda do produto, ou ainda podem ser concedidos indiretamente através do oferecimento de uma ou mais unidades a mais do produto (ou até mesmo de outro produto) de graça no momento da compra. Acredita-se que esses descontos também gerem um aumento significativo das vendas da Kibon.

No entanto, a Kibon não possui um histórico detalhado a respeito das ações realizadas relacionadas à propaganda, promoções e descontos oferecidos aos clientes, com informações sobre o período, o local de ocorrência, a relevância dos mesmos, entre outros.

A única informação disponível relacionada com essas ações é um relatório onde se contabilizam todos os gastos (em unidades monetárias) realizados pelas áreas de marketing e trade marketing, subdivididos em despesas relativas a:

- Material de Merchandising
- Promoções (relativo a quanto foi gasto com as ações promocionais realizadas)
- Divulgação da empresa, marca ou produto na Internet
- *Royalties* (*royalties* pagos a algumas empresas pelo uso de imagens de personagens, por exemplo, nas marcas dos sorvetes)
- Relações Públicas
- Realização de Pesquisa de mercado
- Pagamento à agência de publicidade
- Outros

A utilização das informações deste relatório neste trabalho é inviável, pois os dados existentes são os gastos mensais, num total do Brasil, sem especificar as localidades onde tais ações foram realizadas ou tais despesas foram direcionadas. Além disso, o principal problema é que não está especificado também o produto ou a família de produto ao qual foi destinado cada um desses gastos.

A inexistência das informações relacionadas às ações realizadas por marketing, trade marketing e vendas prejudica o estudo completo proposto por este trabalho, uma vez que acredita-se que estas ações sejam, junto com a temperatura, aquelas que exercem maior impacto no crescimento ou na diminuição das vendas de sorvetes.

Sendo assim, no capítulo 9 será apresentado um procedimento estruturado para a coleta desses dados, para que a empresa passe a guardar um histórico das ações realizadas e, dessa maneira, possa realizar posteriormente estudos análogos aos apresentados neste trabalho, ou sugeridos no capítulo sobre os trabalhos futuros, e obter um maior conhecimento a respeito dos impactos das mesmas.

Dessa maneira, **neste momento só será possível a realização da análise da influência da temperatura sobre o volume de vendas.**

Um instituto de meteorologia disponibiliza para a empresa o histórico de temperaturas médias por dia, por localidade onde a Kibon possui filial. Os dados disponíveis são aqueles desde 2002.

Com relação ao volume de vendas, a empresa possui um banco de dados que guarda este histórico, conforme explicado anteriormente.

Portanto, o histórico de temperatura e o de vendas possuem uma massa de dados significativa, confiável e no nível de detalhe exigido pelos estudos que serão realizados.

### 6.2.3 *Produtos analisados*

Devido à grande variedade de produtos da Kibon, foram escolhidas algumas famílias de produtos específicas para efeito de análise, que serão denominadas por **A, B, C, D e E**. Os verdadeiros nomes das famílias escolhidas foram omitidos a pedido da empresa para se manter a confidencialidade dessas informações.

A escolha de famílias e não SKUs, ou seja, cada item do portfólio da empresa, pode ser explicada pelos seguintes motivos:

- Muitos SKUs entram e saem de linha regularmente e, dessa maneira, não possuem um histórico de dados grande o suficiente para as análises necessárias.
- Dificilmente as ações de vendas e marketing são tomadas sobre um SKU específico; em geral, as promoções e descontos, por exemplo, são dadas sobre uma família inteira de produtos.

Para a definição de quais famílias de produto seriam objetos de análise deste estudo, foi feito uma classificação ABC baseada no princípio do conceito 80-20, ou princípio de Pareto.

De acordo com BALLOU (2001), os produtos de uma empresa contribuem para as vendas e lucros de maneira desproporcional. Isto é, alguns poucos itens podem contribuir com uma proporção elevada para o volume de vendas. Dessa maneira, o conceito 80-20 baseia-se na idéia de que 80% das vendas da empresa são gerados por 20% dos produtos, apesar dessa proporção raramente ser observada na prática.

Feito o levantamento os itens e seus respectivos volumes de vendas, classificam-se os 20% mais importantes como itens A, considerados como os de movimentação rápida; os próximos 30% de itens B (movimentação média) e os restantes de itens C (movimentação lenta).

BALLOU (2001) apresenta a classificação ABC como sendo uma ferramenta para a definição de estratégias de estoque e distribuição. No entanto, para este trabalho, a classificação ABC servirá de ferramenta para identificação das famílias de produtos que seriam as mais importantes para a empresa (itens A) e que, portanto, serão os objetos de análise.

Na figura 6.1, pode-se observar que todos os itens escolhidos para análise estão classificados como itens A quando analisadas todas as famílias de produtos da empresa.

Quando se analisam as famílias separadas em suas respectivas linhas, nas figuras 6.2 e 6.3, é ainda mais clara a importância daquelas escolhidas para a empresa em termos de participação no volume de faturamento total.

Cada número refere-se a outras famílias de produtos que não foram escolhidas para a análise.

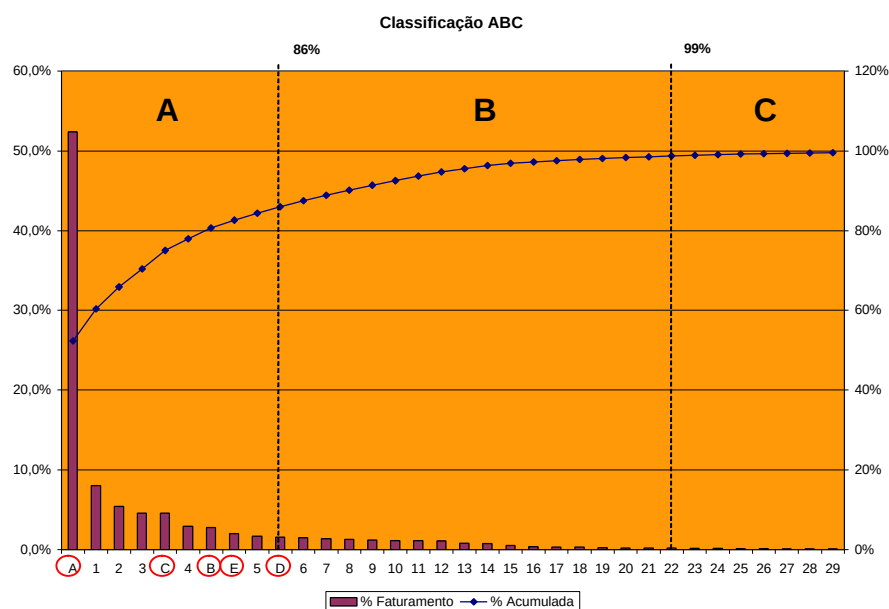


Figura 6.1 - Classificação ABC das famílias de produtos Kibon



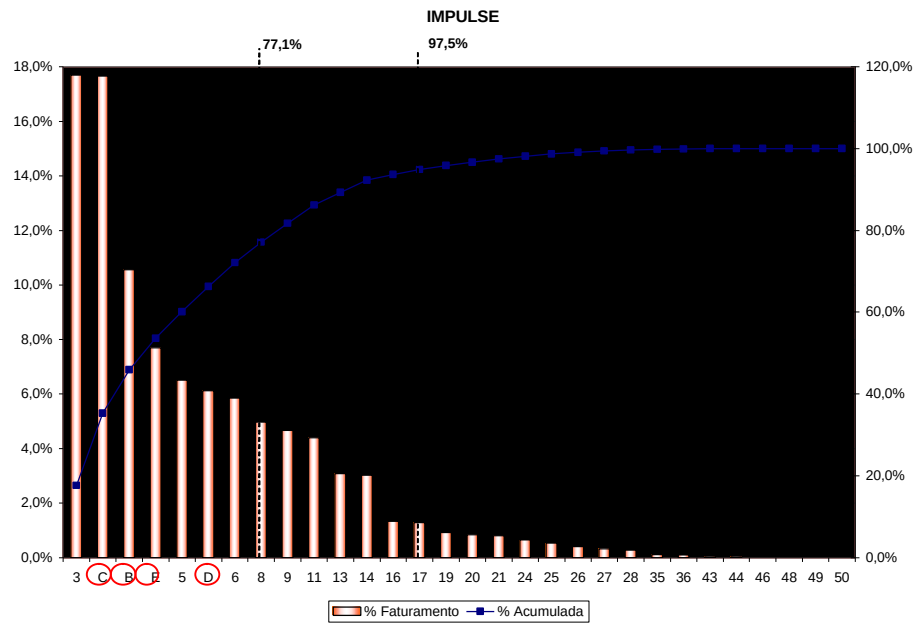


Figura 6.2 - Classificação ABC das famílias da linha Impulse

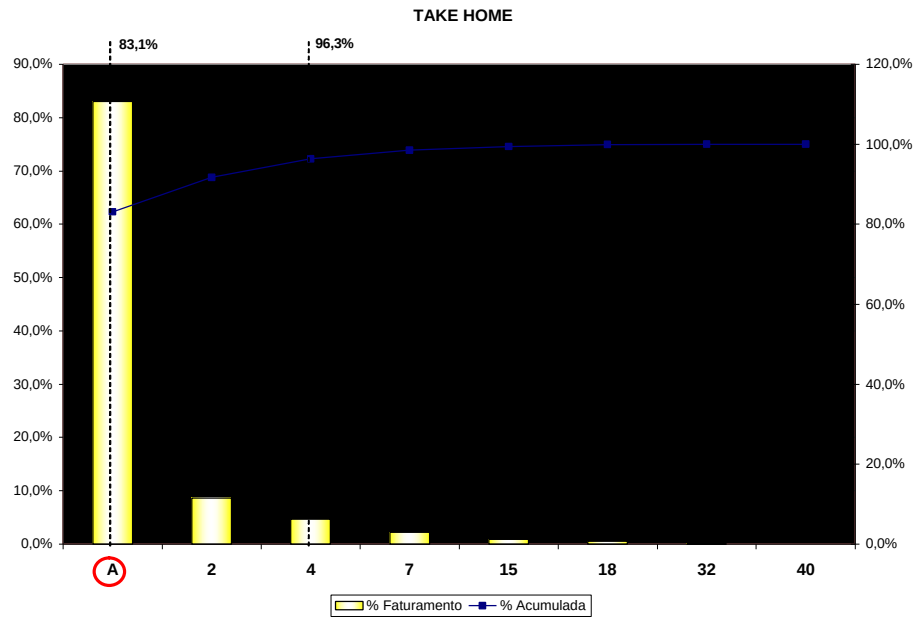


Figura 6.3 - Classificação ABC das famílias da linha Take Home

#### 6.2.4 Filiais escolhidas

Conforme apresentado na introdução deste trabalho, a Kibon possui diversas filiais, que por possuírem comportamentos e volumes diferentes de suas demandas, são abastecidas com diferentes periodicidades: algumas recebem produtos das fábricas diariamente, enquanto outras recebem somente uma vez por semana.

Além disso, algumas ações são realizadas somente em uma ou outra filial específica, e o comportamento do consumidor e da temperatura também varia muito de uma pra outra. Sendo assim, é importante que este estudo seja feito no nível de filial, e não agrupado por total do Brasil ou mesmo por região.

Conforme apresentado no capítulo de apresentação da empresa, a Kibon possui atualmente 15 filiais de vendas.

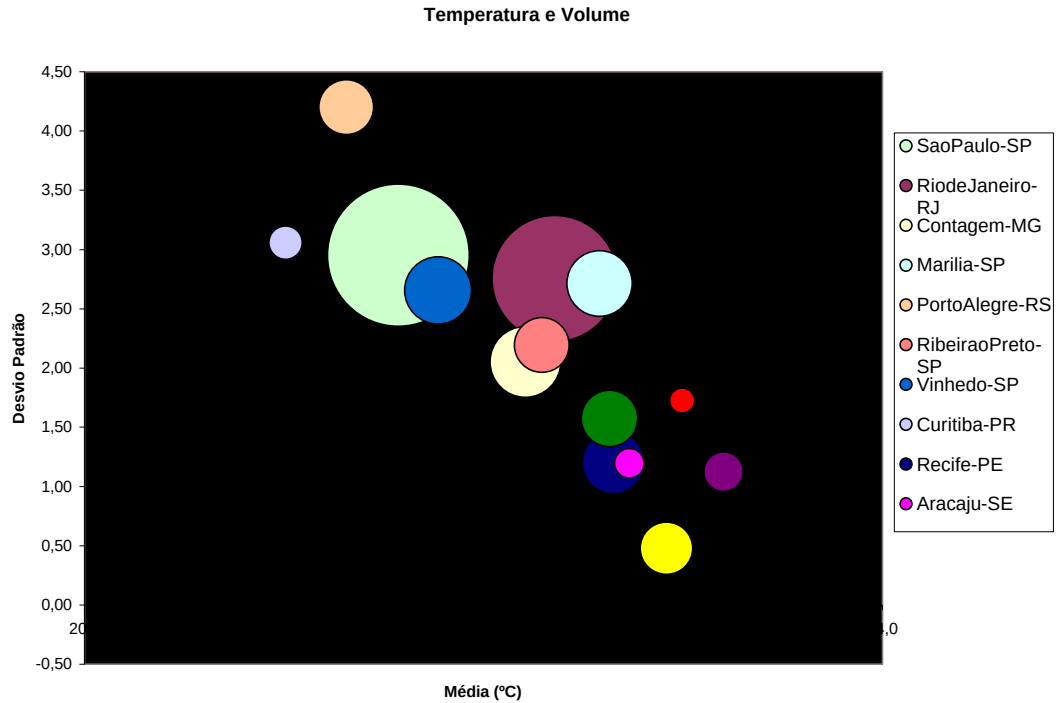
Para a escolha das filiais a serem analisadas, os seguintes aspectos foram levados em consideração:

- Volume de vendas: representa a importância da filial para empresa com relação ao seu faturamento. Foram escolhidas aquelas com maior volume de vendas.
- Temperatura: as temperaturas médias e o desvio padrão de cada localidade das filiais foram levantados.

Através da análise da figura 6.4, percebe-se que as filiais com maior desvio padrão da média das temperaturas são aquelas localizadas na região Sul e Sudeste; em contrapartida, a variação da temperatura ao longo do ano nas filias do Norte e Nordeste é menor, em consequência das altas temperaturas apresentadas, como era de se esperar. Sendo assim, espera-se que nestas últimas a influência da sazonalidade não seja tão acentuada quanto nas filiais do Sul-Sudeste.

**Levando estes aspectos em consideração, foram escolhidas para a realização deste estudo quatro delas: São Paulo, Rio de Janeiro, Recife e Fortaleza.**

As duas primeiras são aquelas que apresentam maior desvio padrão e maior volume de vendas; as duas últimas, são também aquelas de maior volume, mas de menor desvio-padrão.



*Figura 6.4 - Perfil de temperatura e vendas nas filiais*

### 6.3 Período de tempo

O período de tempo a ser analisado será semanal, uma vez que a empresa pretende começar a realizar seu planejamento de produção e o abastecimento de suas filiais tendo em vista esse período de tempo.

Sendo assim, os dados de volume de vendas e temperaturas, disponíveis no formato diário, tiveram que ser consolidados no formato semanal pela própria autora.

**Análise**



## 7. ANÁLISE

Nas próximas seções, serão apresentados os resultados das análises de correlação e autocorrelação dos fatores a serem analisados (temperatura e volume) e da aplicação de alguns modelos de previsão de demanda. Os métodos utilizados foram escolhidos com base no levantamento no capítulo de Revisão Bibliográfica, visando atender aos objetivos deste trabalho.

Para as análises, foram considerados dados semanais desde 2002, o que equivale a uma amostra de  $n=156$  e que, portanto, é adequada para uma análise de correlação, conforme apresentado na seção 5.5.1.

### 7.1 Temperatura (t) x Volume (t)

Nesta primeira análise, busca-se encontrar a relação entre as temperaturas médias e o volume de vendas semanal de cada filial.

Para tanto, foram calculados os valores dos coeficientes de correlação,  $R^2$  e o valor  $P$ , apresentados nas tabelas 7.1, 7.2 e 7.3.

Valores de  $P$  abaixo de 0.05 indicam a existência de uma relação relevante entre as duas variáveis analisadas, conforme explicado na seção 5.5.2. Sendo assim, conclui-se através dos valores apresentados na tabela 7.3 que não se pode afirmar estatisticamente que haja relação entre volume de vendas e temperatura nas seguintes filiais, para as seguintes famílias de produtos:

- Recife: D, cujo valor  $P$  é de 0.229.
- Fortaleza: A, B, C e D, com valores  $P$  respectivos de 0.171, 0.064, 0.616 e 0.470.

Tabela 7.1 - Coeficientes de correlação entre Volume e Temperatura

| Coeficiente de correlação: Volume - Temperatura |      |      |      |      |      |
|-------------------------------------------------|------|------|------|------|------|
|                                                 | A    | B    | C    | D    | E    |
| São Paulo                                       | 0,49 | 0,69 | 0,29 | 0,43 | 0,68 |
| Rio de Janeiro                                  | 0,57 | 0,72 | 0,27 | 0,42 | 0,74 |
| Recife                                          | 0,19 | 0,22 | 0,17 | 0,09 | 0,24 |
| Fortaleza                                       | 0,11 | 0,15 | 0,04 | 0,06 | 0,23 |

Tabela 7.2 -  $R^2$  entre Volume e Temperatura

| R-Quadrado: Volume - Temperatura |       |       |      |       |       |
|----------------------------------|-------|-------|------|-------|-------|
|                                  | A     | B     | C    | D     | E     |
| São Paulo                        | 23,8% | 47,1% | 8,6% | 18,3% | 46,9% |
| Rio de Janeiro                   | 32,0% | 51,5% | 7,3% | 18,0% | 54,0% |
| Recife                           | 3,6%  | 4,9%  | 2,8% | 0,9%  | 5,6%  |
| Fortaleza                        | 1,2%  | 2,4%  | 0,2% | 0,3%  | 5,2%  |

Tabela 7.3 - Valor P da correlação Volume-Temperatura

| P-Value: Volume - Temperatura |       |       |       |       |       |
|-------------------------------|-------|-------|-------|-------|-------|
|                               | A     | B     | C     | D     | E     |
| São Paulo                     | 0,000 | 0,000 | 0,000 | 0,000 | 0,000 |
| Rio de Janeiro                | 0,000 | 0,000 | 0,001 | 0,000 | 0,000 |
| Recife                        | 0,017 | 0,005 | 0,032 | 0,229 | 0,002 |
| Fortaleza                     | 0,171 | 0,064 | 0,616 | 0,470 | 0,005 |

Analisando os dados, nota-se claramente que a correlação entre volume de vendas semanal e temperatura média é maior nas filiais São Paulo e Rio de Janeiro, localizadas na região Sudeste do país, pois estas apresentam maiores valores de  $R^2$ , que representam a porcentagem das vendas que seria explicada pela variação de temperatura.

Os valores para as filiais Recife e Fortaleza são menores, o que representaria uma menor influência da temperatura sobre as vendas, uma vez que elas estão localizadas na região Nordeste do país, onde a variação de temperatura ao longo do ano é pequena, possivelmente gerando uma menor sazonalidade das vendas nessas filiais.

Essa correlação alta ou baixa também pode ser percebida através dos gráficos de pontos do mesmo produto B, mas nas filiais de São Paulo e Recife, visto que, para o caso da primeira, percebe-se mais claramente a existência de uma relação entre as variáveis dos dois eixos (o

aumento de uma é acompanhada de um aumento no valor da outra variável também). Além disso, percebe-se também que, aparentemente, a variabilidade nas vendas aumenta com o aumento da temperatura.

Esses mesmos gráficos indicam a presença de *outliers*, que são alguns pontos significativamente distantes dos demais pra uma mesma faixa de temperatura. Este fato pode já evidenciar uma má qualidade dos dados e pode ter influenciado alguns dos valores encontrados nas tabelas 7.1, 7.2 e 7.3.

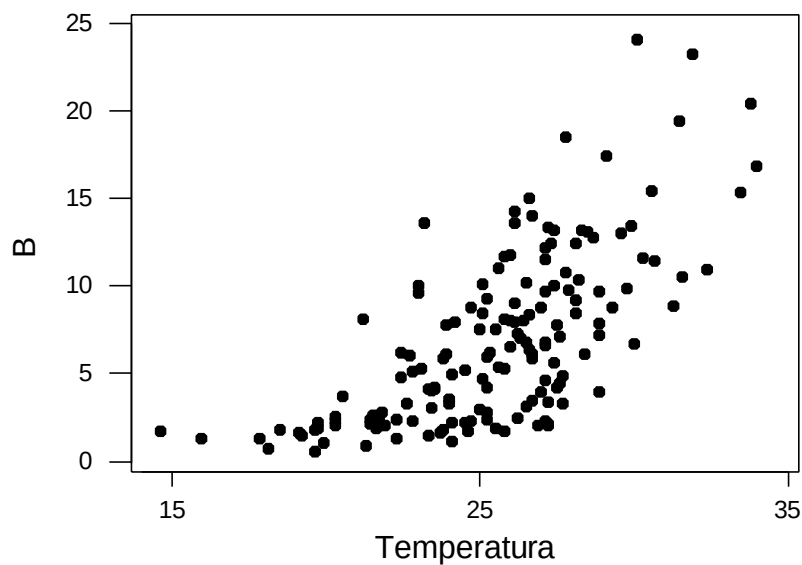


Figura 7.1 - Gráfico Temperatura x Vendas (São Paulo)

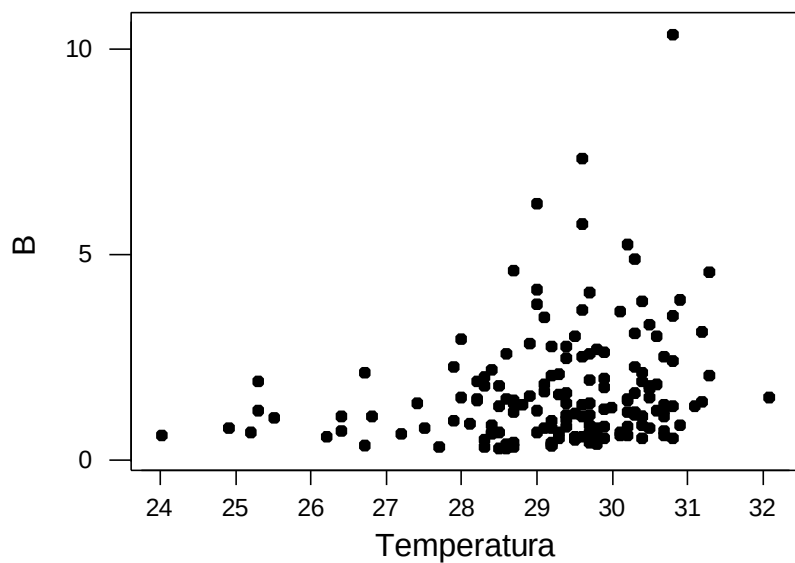


Figura 7.2 - Gráfico Temperatura x Vendas (Recife)

## 7.2 Temperatura (t-1) x Volume (t)

Considerando a hipótese de que o volume de vendas tenha relação não com a temperatura média da própria semana, mas da semana anterior, ou seja, que haja uma defasagem na influência da temperatura sobre um aumento de vendas da Kibon, os coeficientes de correlação foram novamente calculados considerando-se este novo cenário, assim como o valor  $P$ .

Mais uma vez, algumas famílias de produtos em certas filiais apresentaram valores  $P$  acima de 0.05 e, portanto, não se pode afirmar a existência de relação relevante entre o volume de vendas e a temperatura da semana anterior para esses casos. São eles:

- Recife: C e D, com valores  $P$  respectivos de 0.08 e 0.53.
- Fortaleza: A, C e D, com valores  $P$  respectivos de 0.068, 0.237 e 0.367.

Tabela 7.4 - Coeficientes de correlação entre Volume (t) e Temperatura (t-1)

| Coeficiente de correlação: Volume - Temperatura $t-1$ |      |      |      |      |      |
|-------------------------------------------------------|------|------|------|------|------|
|                                                       | A    | B    | C    | D    | E    |
| São Paulo                                             | 0,43 | 0,69 | 0,27 | 0,44 | 0,66 |
| Rio de Janeiro                                        | 0,54 | 0,66 | 0,23 | 0,36 | 0,63 |
| Recife                                                | 0,18 | 0,19 | 0,14 | 0,05 | 0,20 |
| Fortaleza                                             | 0,14 | 0,25 | 0,09 | 0,06 | 0,29 |

Tabela 7.5 -  $R^2$  entre Volume (t) e Temperatura (t-1)

| R-Quadrado: Volume - Temperatura $t-1$ |       |       |      |       |       |
|----------------------------------------|-------|-------|------|-------|-------|
|                                        | A     | B     | C    | D     | E     |
| São Paulo                              | 18,6% | 47,5% | 7,5% | 19,6% | 43,9% |
| Rio de Janeiro                         | 28,6% | 43,4% | 5,3% | 13,1% | 39,4% |
| Recife                                 | 3,1%  | 3,6%  | 1,9% | 0,2%  | 4,2%  |
| Fortaleza                              | 2,0%  | 6,1%  | 0,9% | 0,4%  | 8,7%  |

Tabela 7.6 - Valor  $P$  da correlação Volume (t)-Temperatura (t-1)

| P-Value: Volume - Temperatura $t-1$ |       |       |       |       |       |
|-------------------------------------|-------|-------|-------|-------|-------|
|                                     | A     | B     | C     | D     | E     |
| São Paulo                           | 0,000 | 0,000 | 0,000 | 0,000 | 0,000 |
| Rio de Janeiro                      | 0,000 | 0,000 | 0,004 | 0,000 | 0,000 |
| Recife                              | 0,026 | 0,017 | 0,080 | 0,530 | 0,010 |
| Fortaleza                           | 0,068 | 0,002 | 0,237 | 0,367 | 0,000 |



Comparando os resultados obtidos nas duas abordagens, somente para os casos em que o valor de  $P$  obtido foi menor do que 0.05, nota-se que para algumas filiais e algumas famílias de produtos, o maior  $R^2$  e menor  $P$  são obtidos quando se relaciona o volume de vendas com a temperatura do período anterior. No entanto, quando isto ocorre, os valores são muito próximos ao  $R^2$  e  $P$  da relação Volume (t) e Temperatura (t).

As tabelas 7.7 e 7.8 apresentam os melhores resultados obtidos para cada uma das filiais e famílias e a diferença entre os  $R^2$  das duas análises realizadas.

A análise 1 refere-se aos coeficientes encontrados na seção 7.1 enquanto a análise 2 refere-se à seção 7.2.

*Tabela 7.7 - Maiores  $R^2$  e respectivos Valores  $P$*

|                | A                   | B              | C                   | D              | E              |
|----------------|---------------------|----------------|---------------------|----------------|----------------|
| São Paulo      | 23,8%<br>0,000      | 47,5%<br>0,000 | 8,6%<br>0,000       | 19,6%<br>0,000 | 46,9%<br>0,000 |
| Rio de Janeiro | 32,0%<br>0,000      | 51,5%<br>0,000 | 7,3%<br>0,001       | 18,0%<br>0,000 | 54,0%<br>0,000 |
| Recife         | 3,6%<br>0,017       | 4,9%<br>0,005  | 2,8%<br>0,032       | -              | 5,6%<br>0,002  |
| Fortaleza      | -                   | 6,1%<br>0,002  | -                   | -              | 8,7%<br>0,000  |
|                | $R^2$<br>$P$ -Value | Análise 1      | $R^2$<br>$P$ -Value | Análise 2      |                |

*Tabela 7.8 - Diferença entre os  $R^2$*

|                | A    | B    | C    | D    | E     |
|----------------|------|------|------|------|-------|
| São Paulo      | 5,2% | 0,4% | 1,1% | 1,3% | 3,0%  |
| Rio de Janeiro | 3,4% | 8,1% | 2,0% | 5,0% | 14,6% |
| Recife         | 0,5% | 1,3% | 0,9% |      | 1,5%  |
| Fortaleza      |      | 3,7% |      |      | 3,5%  |
| Média          | 3,9% |      |      |      |       |
| Média          | 2,2% |      |      |      |       |

Sendo assim, de uma maneira geral, considera-se que o volume de vendas semanal da Kibon é mais bem explicado pela temperatura da própria semana em questão, e não pela temperatura da semana anterior, uma vez que os coeficientes de correlação encontrados no primeiro caso são maiores, e os valores  $P$  menores, do que os do segundo. Mesmo quando isto não ocorre, os valores da análise 1 são muito próximos aos da análise 2, conforme os resultados mostrados na tabela 7.8.

### 7.3 Regressão Linear

Uma vez encontrados valores significativos de correlação entre temperatura e volume, de um mesmo período  $t$ , tentou-se aplicar um modelo de regressão linear simples sobre os dados para verificar a validade deste método para a realização de uma previsão de demanda.

Foram realizadas previsões somente para as filiais São Paulo e Rio de Janeiro, uma vez que as demais filiais não apresentaram valores de  $R^2$  superiores a 10%, o que significaria que somente 10% de suas vendas são explicadas pela variação da temperatura. Ou seja, a influência da temperatura é ainda menor do que este valor nas filiais Recife e Fortaleza.

Nas filiais analisadas, todas as famílias de produto escolhidas previamente (A, B, C, D e E) foram consideradas.

O horizonte de previsão definido foi de somente uma semana. Isto porque, na prática, a empresa terá disponível somente a previsão de temperatura da semana seguinte no momento de realização da previsão. Desta maneira, o horizonte de previsão de vendas foi ajustado ao horizonte da temperatura.

Para testar a acuracidade do modelo, 12 previsões foram realizadas, o que equivale a um período de 3 meses. Vale ressaltar que ao se realizar uma nova previsão, os valores reais do período anterior eram adicionados à massa de dados.

Como medida de erro de previsão, considerou-se o erro absoluto percentual, uma vez que é a medida de erro já utilizada pela empresa, na qual é chamada de *Sales Forecast Error (SFE)*.

Com todos esses parâmetros definidos, um modelo de regressão linear simples, relacionando o volume de vendas (variável dependente) à temperatura média (variável independente), foi aplicado aos dados com o auxílio do software MINITAB.

Os resultados, junto com os respectivos valores de  $R^2$  e  $P$ , estão resumidos na tabela 7.9 e na 7.10. Os valores  $P$  encontrados (todos iguais a zero) validam, do ponto de vista estatístico, a aplicação de um regressão linear simples aos dados.

Em destaque, percebe-se que os menores valores de desvio absoluto e maior  $F$  (o que indica a existência de relação entre as variáveis, conforme explicado na seção 5.5.2) correspondem às

famílias de maior  $R^2$ , o que era de se esperar, uma vez que para essas famílias uma maior porcentagem das vendas é explicada pela temperatura.

*Tabela 7.9 - Resultados da Regressão Linear (Filial São Paulo)*

|                     | São Paulo |        |       |       |        |
|---------------------|-----------|--------|-------|-------|--------|
|                     | A         | B      | C     | D     | E      |
| <b>Desvio Médio</b> | 45,4%     | 34,1%  | 51,5% | 45,8% | 24,1%  |
| <b>R-Quadrado</b>   | 23,8%     | 47,1%  | 8,6%  | 18,3% | 46,9%  |
| <b>F Médio</b>      | 45,34     | 131,92 | 11,19 | 25,56 | 129,12 |
| <b>P Médio</b>      | 0,00      | 0,00   | 0,00  | 0,00  | 0,00   |

*Tabela 7.10 - Resultados da Regressão Linear (Filial Rio de Janeiro)*

|                     | Rio de Janeiro |        |       |       |        |
|---------------------|----------------|--------|-------|-------|--------|
|                     | A              | B      | C     | D     | E      |
| <b>Desvio Médio</b> | 25,2%          | 33,7%  | 47,6% | 26,3% | 23,7%  |
| <b>R-Quadrado</b>   | 32,0%          | 51,5%  | 7,3%  | 18,0% | 54,0%  |
| <b>F Médio</b>      | 73,15          | 161,41 | 11,29 | 34,67 | 173,32 |
| <b>P Médio</b>      | 0,00           | 0,00   | 0,00  | 0,00  | 0,00   |

No entanto, os resultados apresentados não são satisfatórios para nenhuma das famílias de produtos em ambas as filiais, uma vez que os valores de desvio absoluto percentual médio encontrado são maiores do que 23%, que está acima da meta estipulada pela empresa que é de 15%.

Sendo assim, a aplicação de um modelo de regressão linear simples, tendo a temperatura como variável de explicação não se mostrou apropriada.

## 7.4 Autocorrelação

Uma hipótese que não foi levantada inicialmente pela empresa foi a de que as vendas de uma semana possam ter relação com as vendas da semana anterior.

Para verificar tal suposição, foi feita uma análise de correlação para verificar a existência da influência das vendas de períodos anteriores sobre as vendas de um certo período  $t$ .

Com o auxílio do software MINITAB, é possível verificar a existência da autocorrelação entre os valores de períodos, ou seja, semanas consecutivas, através dos gráficos das funções de autocorrelação (FAC) e de autocorrelação parcial (FACP), nas quais as linhas horizontais em ambos os lados do eixo servem de guia para identificação dos valores relevantes de cada ordem: são significativos aqueles que ultrapassam os limites definidos pelas mesmas.

Conforme apresentado na seção 5.6, o coeficiente de autocorrelação, mostrado no gráfico da FAC, indicaria o grau de relação entre o valor observado em um certo instante  $t$  e valores de períodos anteriores; já o coeficiente de autocorrelação parcial, indicado no gráfico da FACP, mede a relação entre valores de dois instantes, sem a influência dos demais.

Analisando os gráficos de cada um das filiais e de cada um dos produtos, percebe-se claramente as seguintes características:

- Forte autocorrelação, comprovada pela autocorrelação parcial, de ordem 1, ou seja, grande influência das vendas da semana anterior sobre a semana em questão. As figuras 7.3 e 7.4, relativas à família C na filial São Paulo, exemplificam este caso.
- As filiais do Nordeste apresentam uma autocorrelação menos acentuada do que as filiais do Sudeste, o que pode ser percebido através das figuras 7.5 e 7.6, que representam o comportamento da mesma família C, desta vez na filial Recife.
- A família C, apesar de ter apresentado os menores valores de correlação e  $R^2$  entre volume e temperatura, apresenta os valores mais significativos de autocorrelação. Isto indicaria que suas vendas teriam mais relação com as vendas da semana anterior do que com a temperatura.

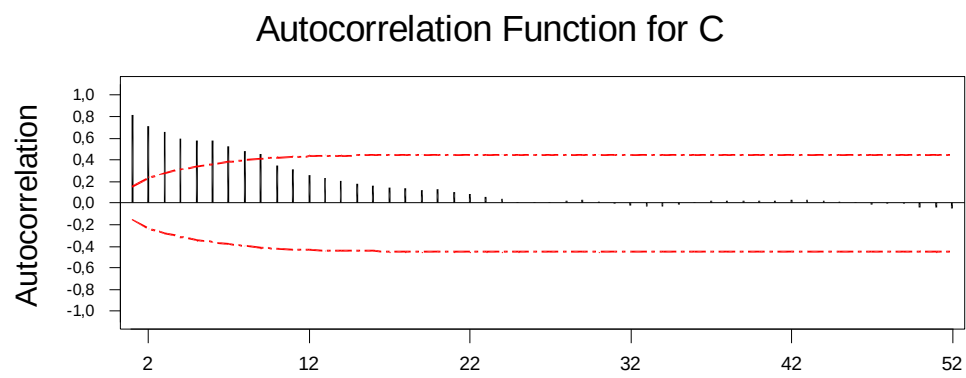


Figura 7.3 - Autocorrelação de C (Filial São Paulo)

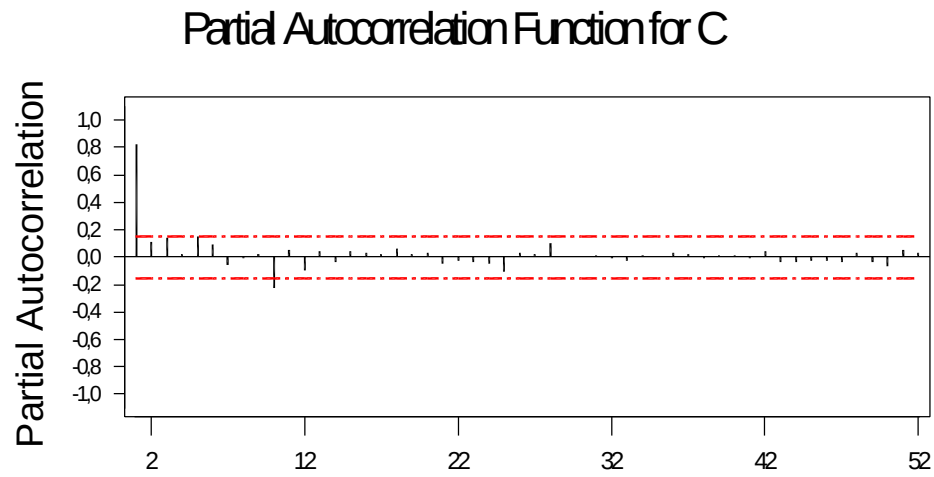


Figura 7.4 - Autocorrelação Parcial de C (Filial São Paulo)

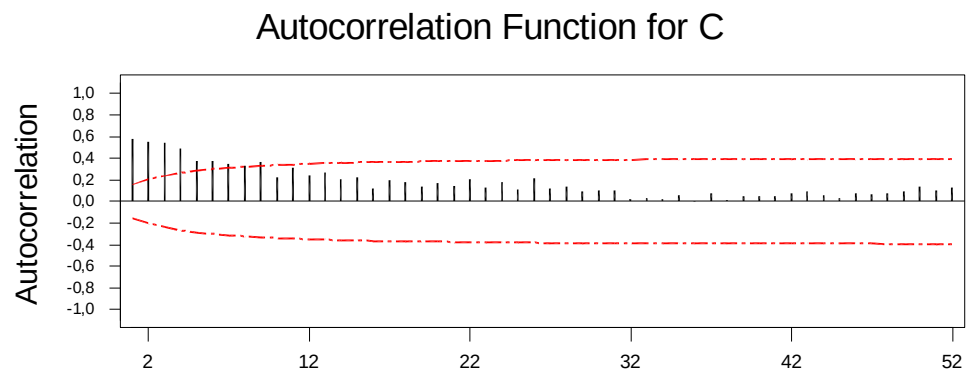
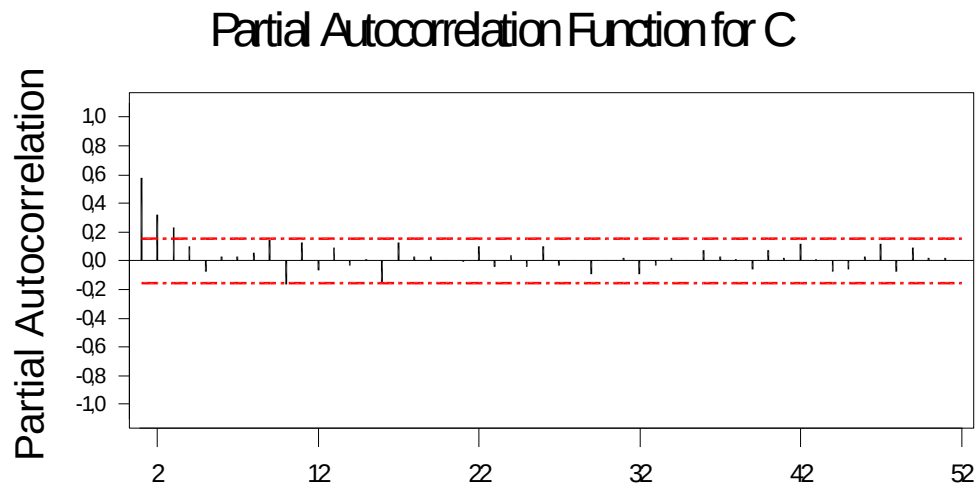


Figura 7.5 - Autocorrelação de C (Filial Recife)



*Figura 7.6 - Autocorrelação Parcial de C (Filial Recife)*

Para efeito de comparação, os valores do coeficiente de autocorrelação parcial de ordem 1 de todas as famílias e filiais estão resumidos na tabela 7.11.

Vale ressaltar que todos os valores  $P$ , para todas as famílias em todas as filiais, foram iguais a zero, o que confirma a existência de uma autocorrelação significativa de ordem 1.

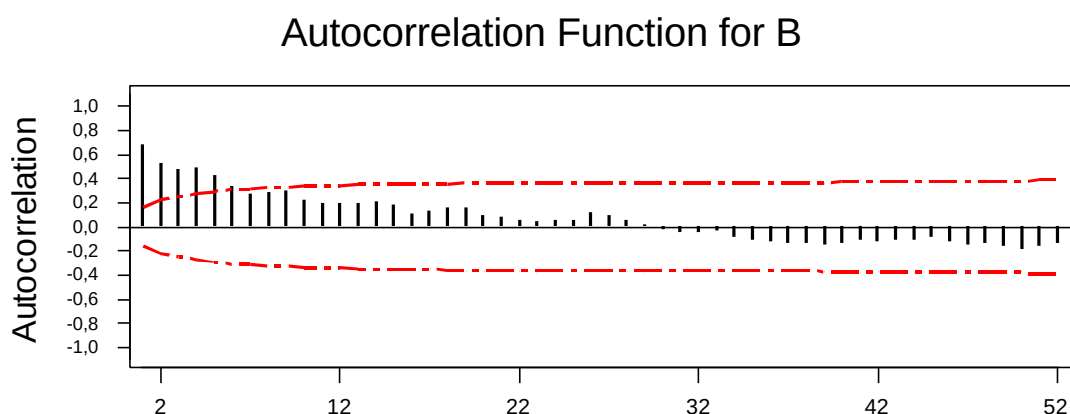
*Tabela 7.11 - Autocorrelação Parcial de ordem 1 por Filial e Família*

| <b>Auto-correlação parcial - Ordem 1</b> |          |          |          |          |          |              |
|------------------------------------------|----------|----------|----------|----------|----------|--------------|
|                                          | <b>A</b> | <b>B</b> | <b>C</b> | <b>D</b> | <b>E</b> | <b>Média</b> |
| <b>São Paulo</b>                         | 0,59     | 0,67     | 0,82     | 0,67     | 0,64     | 0,68         |
| <b>Rio de Janeiro</b>                    | 0,71     | 0,74     | 0,86     | 0,86     | 0,68     | 0,77         |
| <b>Recife</b>                            | 0,31     | 0,37     | 0,58     | 0,30     | 0,38     | 0,39         |
| <b>Fortaleza</b>                         | 0,38     | 0,68     | 0,73     | 0,59     | 0,59     | 0,59         |
| <b>Média</b>                             | 0,50     | 0,62     | 0,75     | 0,61     | 0,57     |              |

## 7.5 Modelo ARIMA

Verificada a existência de uma autocorrelação entre os volumes de vendas, é preciso identificar qual o modelo ARIMA é mais apropriado para representar a série.

Em primeiro lugar, vale ressaltar que não será considerado um modelo sazonal em nenhum dos casos estudados. De acordo com MAKRIDAKIS (1998), o modelo sazonal é considerado quando há uma autocorrelação significativa nas ordens equivalentes a duração do período de sazonalidade, ou seja, à ordem 52 nos casos em questão. No entanto, em nenhuma das filiais e famílias analisadas este fato é constatado, como no exemplo da figura 7.7, que mostra a função FAC na filial Recife, para a família B. Note que o valor de autocorrelação para esta ordem é baixo e sua respectiva barra não ultrapassa os limites definidos pelas duas linhas horizontais.



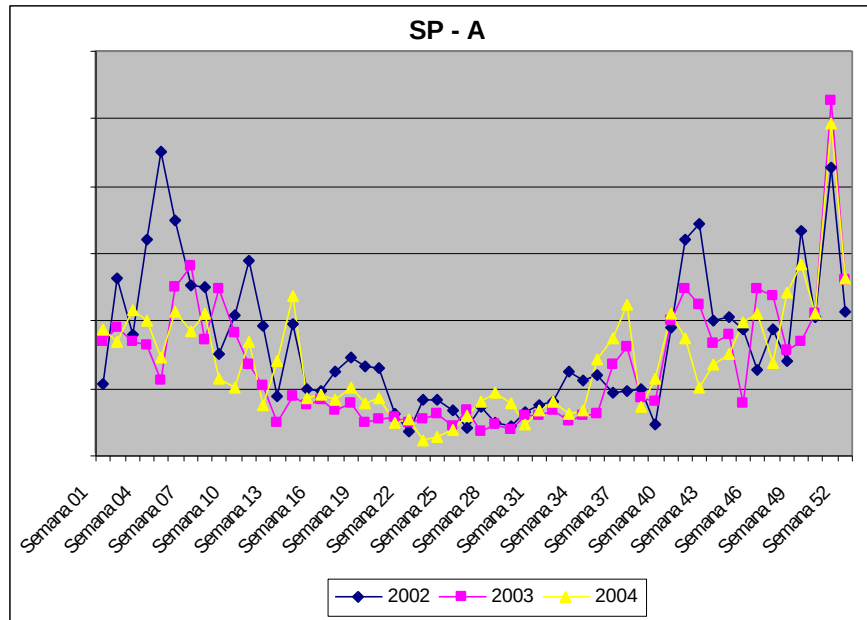
*Figura 7.7 - Função de autocorrelação até a ordem 52*

Seguindo os passos apresentados na seção 5.6, foram construídos gráficos para verificar se a série em questão é do tipo estacionária.

Para todas as filiais e produtos, não foi constatada a presença de uma tendência de crescimento ou diminuição das vendas nos últimos três anos, como no exemplo da figura 7.8, que mostra as vendas da família A, na filial São Paulo. Note que as linhas praticamente sobrepostas indicam que as vendas da empresa vêm se mantendo sem mudanças significativas, para os mesmos períodos, nos últimos três anos, o que é constatado também para as demais famílias analisadas, em todas as filiais. Sendo assim, pode-se considerar todas



as séries como sendo estacionárias e, portanto, não foi necessária a realização de uma diferenciação, de modo que todos os modelos serão do tipo  $(p,0,q)$ .



*Figura 7.8 - Vendas da família A na Filial São Paulo*

Para definir os demais parâmetros do modelo ARIMA, foi feita uma análise das funções de autocorrelação (FAC) e de autocorrelação parcial (FACP), conforme apresentado por RAMOS (1999) e na seção 5.6.

Com base no comportamento das funções FAC e FACP, até dois modelos ARIMA foram escolhidos para a análise. Em alguns casos específicos, que serão detalhados mais adiante, mais de dois modelos tiveram que ser testados.

Definidos os modelos que seriam mais apropriados para cada filial e produto, os seguintes resultados foram analisados para se concluir a adequação ou não dos modelos:

- Valores previstos e o erro absoluto percentual médio: esses indicam a aplicabilidade do modelo para a previsão de demanda e o grau de acuracidade proporcionado.
- Função de autocorrelação e autocorrelação parcial dos resíduos: não deve ser identificada correlação significativa entre os resíduos dos diversos períodos, o que indicaria uma não adequação do modelo aplicado.

- Valor  $P$  para cada um dos coeficientes do modelo: o valor  $P$  deve apresentar valores baixos, em geral menores do que 0.05, para que o coeficiente possa ser considerado significativo, conforme já explicado anteriormente.
- Gráfico dos resíduos x variável independente, no caso o próprio volume de venda do período. Mais uma vez, a existência de uma relação entre os resíduos e a variável indicaria a não adequação do modelo.
- Normalidade dos resíduos, que devem apresentar distribuição normal.

Os casos a seguir que merecem um maior destaque:

### **Recife – Família A**

Os gráficos de autocorrelação e autocorrelação parcial apresentam neste caso um comportamento que dificulta a identificação do modelo ARIMA mais adequado: a função FAC diminui gradualmente, mas oscila para cima e para baixo até a ordem 9, quando deixa de apresentar valores significativos; já a função FACP apresenta valores relevantes até a ordem 4, mas seu valor não é relevante para a ordem 3.

Com base nesta primeira análise não há como se definir um ou dois modelos que seriam os mais adequados e, sendo assim, 4 modelos foram testados: (4,0,0), (3,0,0), (4,0,4) e (4,0,3). A escolha desses modelos foi feita com base nas características das funções FAC e FACP que foram possíveis de serem identificadas.

O modelo (4,0,4), foi descartado uma vez que apresentou um valor  $P$  de 0,082 para o coeficiente de média móvel da ordem 4, o que indicaria que esse coeficiente não é relevante no modelo.

Já o segundo deles, (3,0,0), não se mostrou adequado para os dados em questão, uma vez que os gráficos FAC e FACP dos resíduos mostram a existência de uma correlação entre os mesmos.

Por fim, comparando os modelos (4,0,3) e (4,0,0), tanto os gráficos FAC e FACP quanto o de teste normal dos resíduos, apresentam melhores resultados para o modelo (4,0,3), o que explica a escolha deste como sendo aquele que melhor se ajusta aos dados em questão.

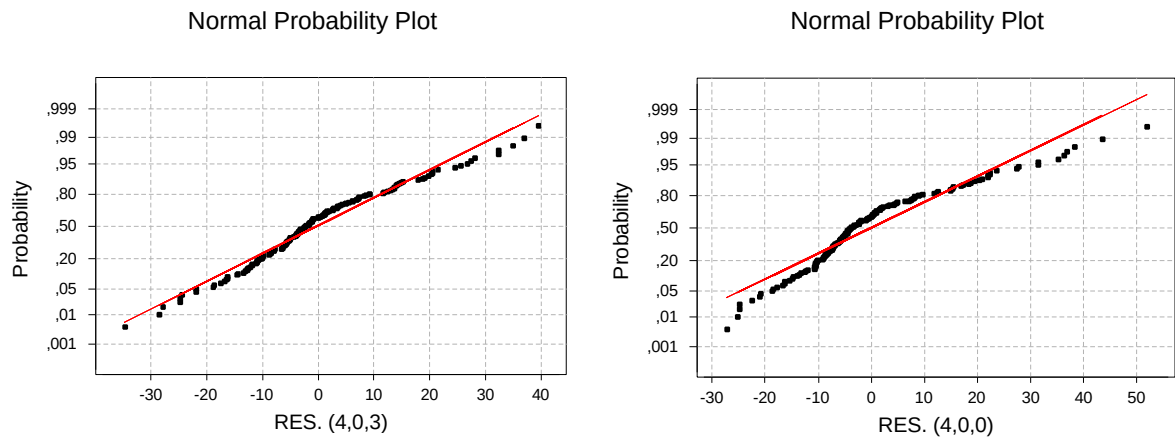


Figura 7.9 - Comparação entre o teste normal dos modelos  $(4,0,3)$  e  $(4,0,0)$

Através da análise das figuras 7.10 e 7.11, percebe-se que não há correlação entre os resíduos de diferentes períodos de tempo, o que comprova a adequação do modelo pra representar a série em questão.

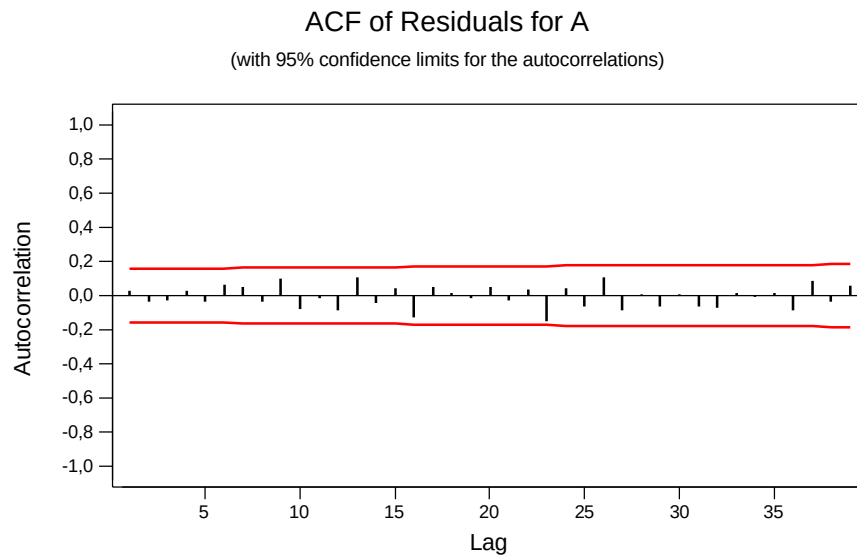
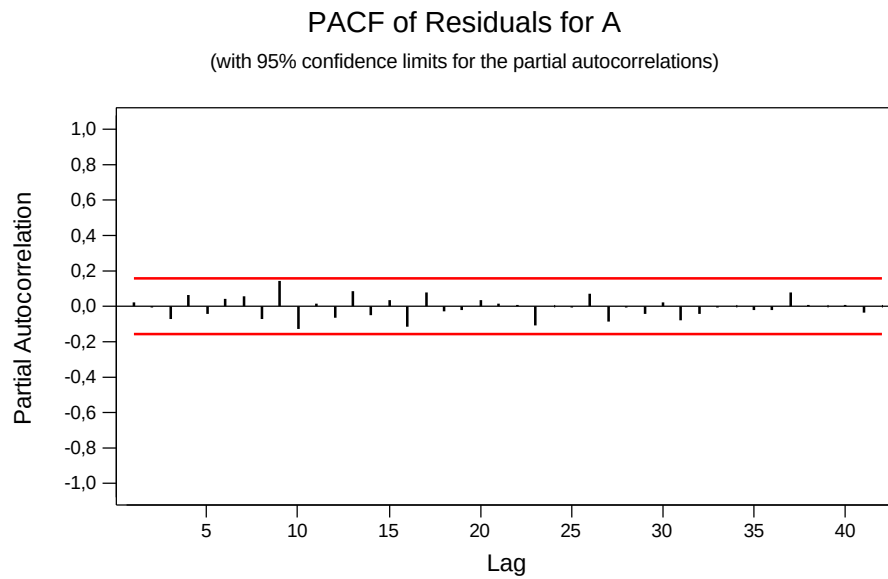


Figura 7.10 - Função FAC do modelo  $(4,0,3)$  para A na Filial Recife



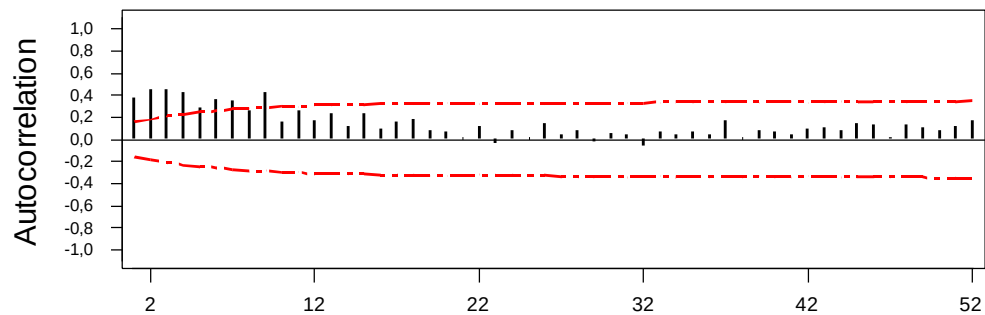
*Figura 7.11 - Função FACP do modelo (4,0,3) para A na Filial Recife*

### **Recife – Família E**

Para este caso, também foram testados quatro modelos, com base no comportamento da função FAC (diminui gradualmente, mas oscila para cima e para baixo até a ordem 9) e da FACP (apresenta valores relevantes até a ordem 3, mas os valores para as ordens 1 e 2 são mais significativos), conforme apresentado nas figuras 7.12 e 7.13.

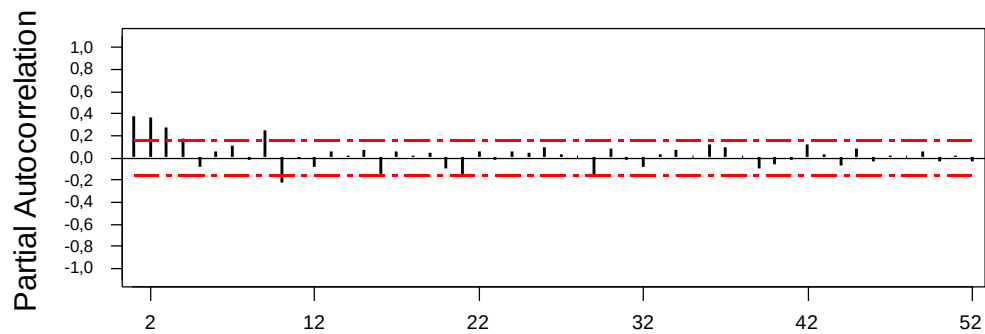
Os modelos cujo coeficiente de Média Móvel é igual a zero não apresentaram resultados satisfatórios, uma vez que foi constatada uma correlação entre os resíduos, o que levou o modelo (3,0,3) a ser escolhido como o mais adequado para os dados em questão.

### Autocorrelation Function for E



*Figura 7.12 - FAC para Família E (Filial Recife)*

### Partial Autocorrelation Function for E



*Figura 7.13 - FACP para Família E (Filial Recife)*

Os modelos selecionados para cada análise em cada uma das filiais, para cada família de produto, estão resumidos na tabela 7.12. Estão destacados aqueles modelos que se mostraram mais adequados para representar a série com base na análise dos resultados das funções FAC e FACP, além de outros gráficos já citados anteriormente.

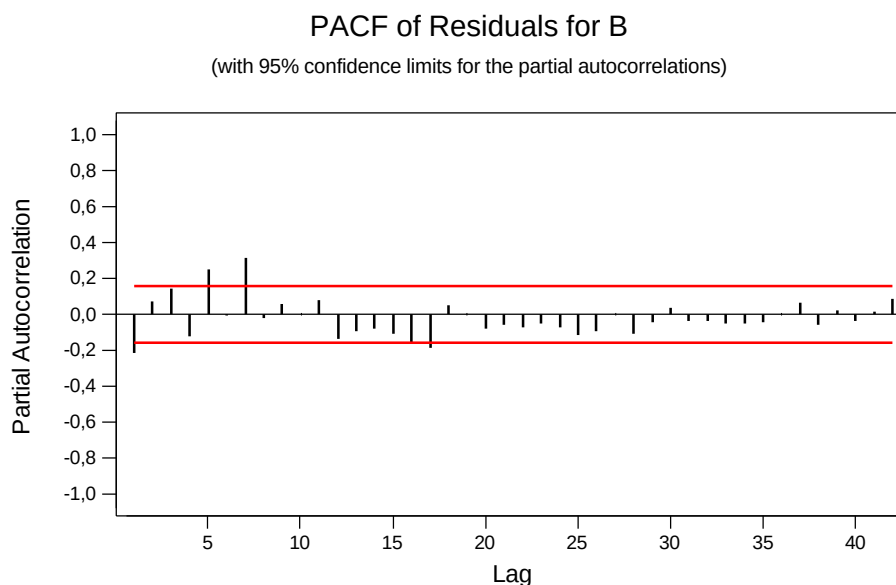
Tabela 7.12 - Modelos ARIMA escolhidos

| Filial         | Família | FAC                  | FACP                                                                                                                     | MODELO 1 | MODELO 2 | MODELO 3 | MODELO 4 |
|----------------|---------|----------------------|--------------------------------------------------------------------------------------------------------------------------|----------|----------|----------|----------|
| São Paulo      | A       | Diminui gradualmente | Se anula após a ordem 2, mas a autocorrelação parcial é muito mais significativa para a ordem 1                          | (1,0,0)  | -        | -        | -        |
| São Paulo      | B       | Diminui gradualmente | Se anula após a ordem 2, mas a autocorrelação parcial é muito mais significativa para a ordem 1                          | (1,0,0)  | -        | -        | -        |
| São Paulo      | C       | Diminui gradualmente | Se anula após a ordem 1                                                                                                  | (1,0,0)  | -        | -        | -        |
| São Paulo      | D       | Diminui gradualmente | Se anula após a ordem 1                                                                                                  | (1,0,0)  | -        | -        | -        |
| São Paulo      | E       | Diminui gradualmente | Se anula após a ordem 1                                                                                                  | (1,0,0)  | -        | -        | -        |
| Rio de Janeiro | A       | Diminui gradualmente | Se anula após a ordem 2, mas a autocorrelação parcial é muito mais significativa para a ordem 1                          | (1,0,0)  | -        | -        | -        |
| Rio de Janeiro | B       | Diminui gradualmente | Se anula após ordem 5, mas não há correlação significativa para ordens 3 e 4                                             | (1,0,0)  | (2,0,0)  | -        | -        |
| Rio de Janeiro | C       | Diminui gradualmente | Se anula após a ordem 1                                                                                                  | (1,0,0)  | -        | -        | -        |
| Rio de Janeiro | D       | Diminui gradualmente | Se anula após ordem 3, mas não há correlação significativa para ordem 2 e a correlação para ordem 1 é mais significativa | (1,0,0)  | -        | -        | -        |
| Rio de Janeiro | E       | Diminui gradualmente | Se anula após a ordem 2                                                                                                  | (1,0,0)  | (2,0,0)  | -        | -        |

Tabela 7.12 - Continuação

| Filial    | Família | FAC                                                  | FACP                                                                                            | MODELO 1 | MODELO 2 | MODELO 3 | MODELO 4 |
|-----------|---------|------------------------------------------------------|-------------------------------------------------------------------------------------------------|----------|----------|----------|----------|
| Recife    | A       | Diminui gradualmente, mas oscila nas diversas ordens | Se anula após ordem 4, mas não há correlação significativa para ordem 3                         | (4,0,0)  | (3,0,0)  | (4,0,4)  | (4,0,3)  |
| Recife    | B       | Diminui gradualmente                                 | Se anula após a ordem 3, mas os coeficientes para as ordens 1 e 2 são mais significativos       | (1,0,0)  | (2,0,0)  | -        | -        |
| Recife    | C       | Diminui gradualmente                                 | Se anula após a ordem 3, mas os coeficientes para as ordens 1 e 2 são mais significativos       | (1,0,0)  | (2,0,0)  | -        | -        |
| Recife    | D       | Significativo somente para ordem 1                   | Significativo somente para ordem 1                                                              | (1,0,0)  | (1,0,1)  | -        | -        |
| Recife    | E       | Diminui gradualmente, mas oscila nas diversas ordens | Se anula após a ordem 3, mas os coeficientes para as ordens 1 e 2 são mais significativos       | (3,0,0)  | (2,0,0)  | (1,0,0)  | (3,0,3)  |
| Fortaleza | A       | Diminui gradualmente, mas oscila nas diversas ordens | Se anula após ordem 4, mas não há correlação significativa para ordem 3 e 2                     | (4,0,0)  | (1,0,0)  | (4,0,4)  | -        |
| Fortaleza | B       | Diminui gradualmente                                 | Se anula após a ordem 1                                                                         | (1,0,0)  | -        | -        | -        |
| Fortaleza | C       | Diminui gradualmente                                 | Se anula após a ordem 2, mas a autocorrelação parcial é muito mais significativa para a ordem 1 | (1,0,0)  | -        | -        | -        |
| Fortaleza | D       | Diminui gradualmente                                 | Se anula após a ordem 1                                                                         | (1,0,0)  | -        | -        | -        |
| Fortaleza | E       | Diminui gradualmente                                 | Se anula após a ordem 2, mas a autocorrelação parcial é muito mais significativa para a ordem 1 | (1,0,0)  | -        | -        | -        |

Após a análise de todas as filiais e de todas as famílias, constatou-se que somente nas famílias B, C e E, na filial Rio de Janeiro, o modelo escolhido não se mostrou adequado já na análise das funções FAC e FACP dos resíduos, uma vez que as funções indicaram uma autocorrelação significativa entre os resíduos nas ordens 1, 5 e 6, conforme exemplo da figura 7.14.



*Figura 7.14 - Função FACP dos resíduos da Família B (Filial Rio de Janeiro)*

Para todas as outras famílias de produto e filiais, foram geradas 12 previsões com base no modelo mais adequado definido para cada caso. Os valores médios de desvio absoluto percentual médio encontrados estão resumidos na tabela 7.13.

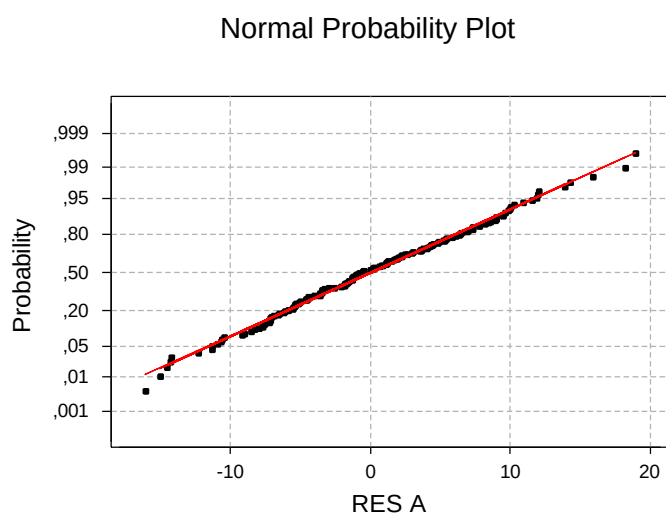
*Tabela 7.13 – Erro Absoluto Percentual Médio (ARIMA)*

| <b>Erro absoluto percentual Médio - ARIMA</b> |          |          |          |          |          |              |
|-----------------------------------------------|----------|----------|----------|----------|----------|--------------|
|                                               | <b>A</b> | <b>B</b> | <b>C</b> | <b>D</b> | <b>E</b> | <b>Média</b> |
| <b>São Paulo</b>                              | 51%      | 47%      | 38%      | 32%      | 36%      | 41%          |
| <b>Rio de Janeiro</b>                         | 11%      |          |          | 11%      |          | 11%          |
| <b>Recife</b>                                 | 26%      | 14%      | 23%      | 17%      | 16%      | 19%          |
| <b>Fortaleza</b>                              | 31%      | 14%      | 19%      | 12%      | 14%      | 18%          |
| <b>Média</b>                                  | 30%      | 25%      | 27%      | 18%      | 22%      |              |

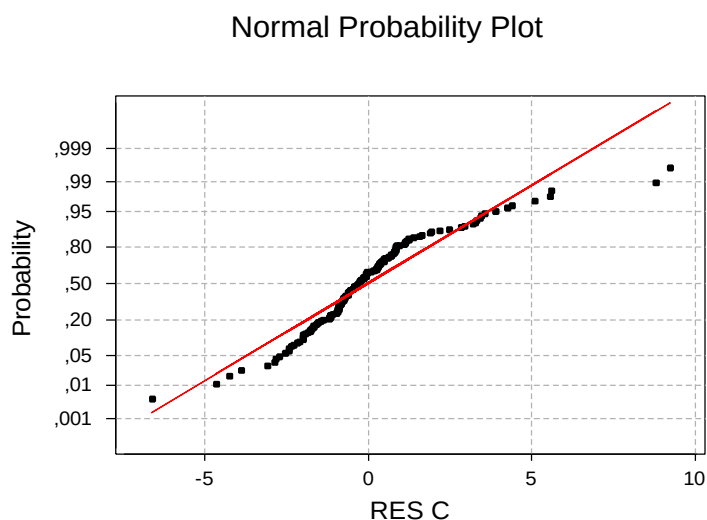


Os resultados encontrados são satisfatórios em grande parte dos casos analisados, uma vez que o erro absoluto percentual médio apresenta valor entre 11% e 20%, como no caso da filial Rio de Janeiro.

No entanto, ao se analisar os resíduos mais profundamente, nota-se no gráfico do teste normal que, na maior parte dos casos estudados, os resíduos não apresentam distribuição normal. De fato, somente os resíduos do modelo aplicado para a família A na filial Fortaleza apresentam distribuição normal, conforme mostrado na figura 7.15, na qual percebe-se que os pontos estão ajustados à reta. Este fato não ocorre nos demais casos, conforme exemplo da figura 7.16.



*Figura 7.15 - Gráfico de probabilidade normal para Família A (Filial Fortaleza)*

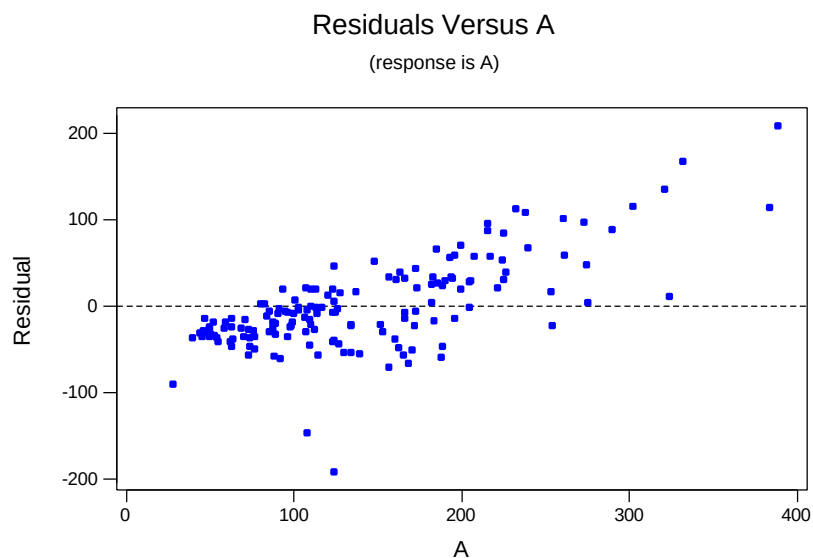


*Figura 7.16 - Gráfico de probabilidade normal para Família C (Filial São Paulo)*

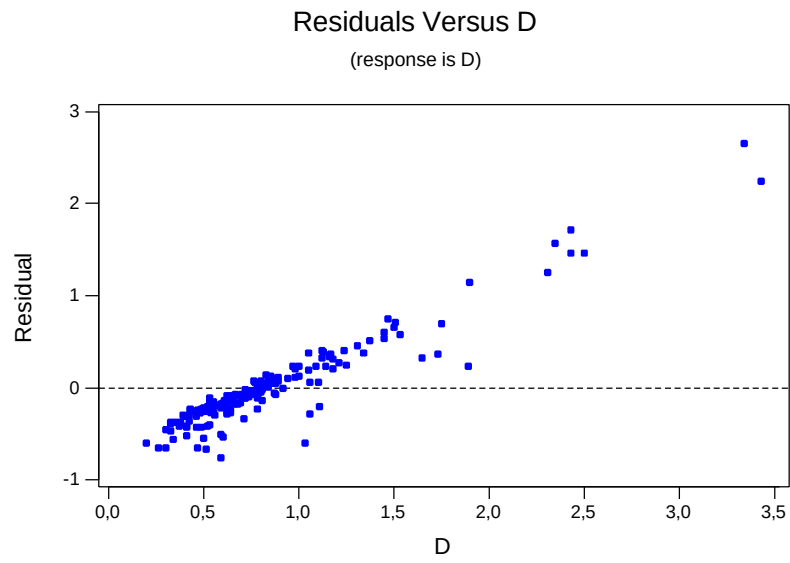
Além disso, o gráfico dos resíduos em função da variável independente, exemplificado nas figuras 7.17 e 7.18, mostra que há relação entre os mesmos e a variável, no caso o volume de vendas, uma vez que os pontos não estão distribuídos aleatoriamente nos dois lados do eixo horizontal, apresentando irregularidades como:

- Tendência dos pontos (aumento dos resíduos com o aumento no valor da variável)
- Pontos consecutivos do mesmo lado do eixo
- Pontos muito distantes dos demais pontos

Este fato comprova que o ajuste do modelo não foi satisfatório, e, conseqüentemente, invalidaria os resultados obtidos na realização da previsão. Além disso, vale ressaltar que isto pode ser um indício da má qualidade dos dados históricos, o que impossibilitaria o atingimento de um resultado satisfatório, quaisquer que fosse o modelo testado.



*Figura 7.17 - Gráfico de resíduos da Família A (Filial Rio de Janeiro)*



*Figura 7.18 - Gráfico de resíduos da Família D (Filial Recife)*

## Resultados

## 8. RESULTADOS

**A existência de uma relação entre o volume de vendas da filial e a temperatura média registrada foi verificada somente nos casos da filial São Paulo e Rio de Janeiro, o que leva a concluir que esse comportamento deve ser verificado também nas demais filiais da região Sul-Sudeste, uma vez que o desvio padrão da temperatura nessas localidades também é alto.**

Por outro lado, não há relação significativa entre as vendas e a temperatura nas filiais localizadas no Nordeste, o que comprova a hipótese levantada no momento de escolha das filiais a serem analisadas de que a relação entre venda e temperatura não é tão acentuada nesta região do país, onde as temperaturas são altas e com pouca variação ao longo do ano.

Os respectivos coeficientes de correlação e  $R^2$  encontrados, assim como o valor  $P$ , nas filiais São Paulo e Rio de Janeiro, comprovam que o volume de vendas dessas localidades é explicada em até 54% (como no caso da família E no Rio de Janeiro) pela variação da temperatura.

No entanto, os resultados não satisfatórios obtidos com a aplicação de um modelo de regressão linear, relacionando as vendas com a temperatura, indicam que este não é o único fator que exerceria impacto sobre as vendas e, portanto, a regressão linear simples não se mostrou adequada para as séries.

Apesar de não ter sido mencionado pela empresa, através da análise de autocorrelação, comprovou-se que **existe relação significativa entre os volumes de vendas de uma semana com os das semanas anteriores.**

No entanto, constatou-se, através dos resíduos obtidos, que não é adequada a aplicação dos modelos ARIMA escolhidos para cada caso. Porém, este fato não pode ser comprovado, uma vez que o problema pode estar na má qualidade dos dados históricos guardados pela empresa.

Sendo assim, fica claro que estudos como este em questão, que buscam relacionar algumas variáveis com o volume de vendas da empresa, ou mesmo a aplicação de outros modelos de previsão de demanda ficam prejudicados devido à falta de dados ou à má qualidade dos mesmos.

Com base nesta constatação, será apresentado no capítulo 9 um procedimento para a coleta de dados, visando a geração de um banco de dados de qualidade e contendo todas as informações necessárias para análises futuras mais aprofundadas no tema.

**Trabalhos futuros**

A thick, horizontal orange brushstroke underline that starts with a light, faded edge on the left and becomes a solid, vibrant orange towards the right, ending in a slightly irregular, hand-painted style.

## 9. TRABALHOS FUTUROS

### 9.1 Procedimento para coleta de dados

Através do estudo realizado e apresentado nas seções anteriores, conclui-se que há a necessidade de uma melhor coleta de dados por parte da empresa.

A qualidade desta coleta não está relacionada somente com a precisão dos dados e confiabilidade dos mesmos, mas também à escolha correta das informações que se deseja analisar e que, portanto, devem ser coletadas e armazenadas.

Para tanto, é necessário em primeiro lugar definir corretamente dois pontos: quais seriam as informações mais importantes a serem levantadas e os atributos relacionados a elas.

OLIVEIRA (1994) apresenta algumas informações que são necessárias para o controle e avaliação da estrutura organizacional, rotinas e procedimentos de uma empresa. São elas:

- Data da ocorrência dos eventos
- Interação entre os dados
- Frequência da coleta dos dados
- Fontes e fluxo de informação
- Possível codificação
- Nível de detalhamento (neste caso, com relação ao produto, tempo e região analisados)
- Pessoas envolvidas / responsáveis
- Áreas responsáveis / envolvidas

Com base nessas informações e tendo em vista que o nível de detalhamento dos dados será igual àquele apresentado neste estudo, ou seja, **por filial, família de produto e por semana**, os possíveis fatores a serem levantados, divididos em fatores internos e externos, serão apresentados na seção 9.1.1.



### 9.1.1 Fatores e atributos

#### Fatores externos

Fatores externos são aqueles que não estão sobre o controle da empresa, mas que podem influenciar as vendas da mesma. São eles:

**Clima:** sendo o sorvete um produto altamente sazonal, tendo as suas maiores vendas nos períodos quentes do ano, acredita-se que o clima seja um dos fatores externos mais impactantes para as vendas da Kibon. Pode ser dividido em:

- **Temperatura máxima, média e mínima:** o presente estudo foi realizado considerando-se somente a temperatura média. No entanto, as vendas da empresa podem ser melhor explicadas, ou seja, podem ter uma maior relação, com a temperatura máxima ou mínima, o que torna interessante o levantamento dessas informações, que poderão ser utilizadas posteriormente em um estudo análogo ao que foi desenvolvido na seção 7.1.
- **Dias de sol:** a relação das vendas com o clima pode ser explicada não pela temperatura, mas sim pela quantidade de dias de sol constatadas na semana. Para se levantar essa informação, a empresa deverá estabelecer um critério claro e bem definido para o que será considerado como sendo um dia de sol. O instituto de meteorologia, com o qual a empresa já tem parceria, poderia ser contatado para ajudar nesta definição.

#### Fatores internos

Ao contrário dos fatores externos, os internos correspondem àqueles que estão sob o controle da empresa ou àqueles que resultam das ações tomadas pela mesma.

Neste grupo encontram-se também todas as ações de marketing e vendas cujos impactos sobre a venda de determinado produto, ou sobre todos os produtos, acredita-se que seja grande, conforme já citado no capítulo de levantamento de dados.

Dessa maneira, pode-se subdividir os fatores internos da seguinte maneira:

**Stock Out:** o chamado *Stock Out* ocorre quando se constata a falta total, ou quase total, do produto em uma determinada filial. Isso pode gerar um volume de vendas que não corresponde à real demanda do mercado, uma vez que esta última pode ter sido maior do que a oferta disponível do produto, e, conseqüentemente, o dado registrado apresentará um comportamento incomum em relação à série.

Essa informação, a ocorrência ou não de *Stock Out* de um produto em uma determinada filial, já é levantada semanalmente pela área de Planejamento da empresa. Sendo assim, há somente a necessidade de se passar a guardar um histórico dessas ocorrências.

**Ações de Marketing:** deve haver um registro de todas as ações de marketing realizadas pela empresa, tais como anúncios em televisão e em outros meios, como revistas, Internet, e outdoors. Cada tipo de anúncio pode ser considerado como sendo um fator diferente, assim como as outras ações realizadas pela área de marketing com o objetivo de aumentar a exposição do produto ou da marca.

**Ações de Trade Marketing e Vendas:** neste grupo se enquadram todas as promoções realizadas e descontos oferecidos aos clientes. Assim como no caso anterior, cada um dos tipos de desconto e promoção podem ser analisados como sendo fatores diferentes.

Analogamente à tabela de estrutura de dados apresentado por KIYUZATO (2001), a figura 9.1 mostra o relacionamento entre estes diversos fatores e o nível de detalhe determinado (filial, semana e família de produto).

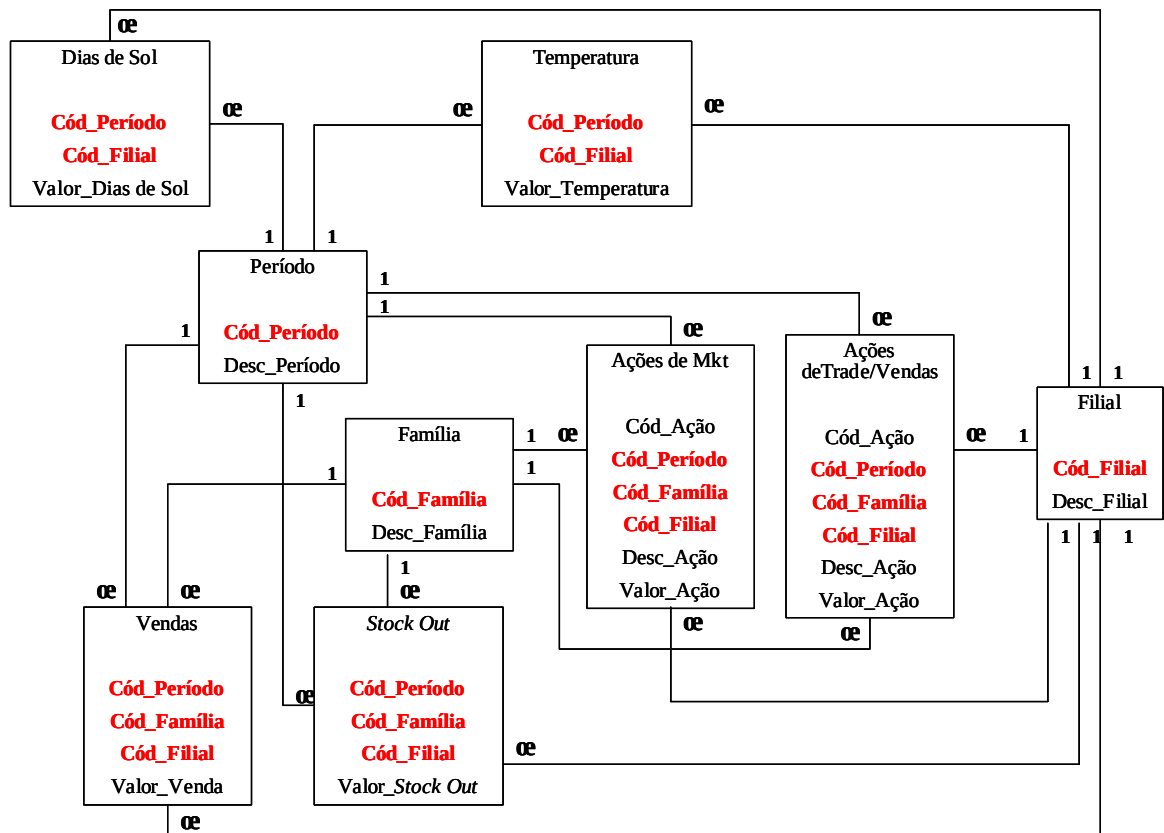


Figura 9.1 - Tabelas para coleta de dados

Elaborado pela autora

A tabela 9.1 resume a sugestão de todos os fatores e respectivos atributos que devem ser levantados, assim como suas unidades de coleta.

Tabela 9.1 - Resumos dos fatores a serem levantados

| 1     | FATORES EXTERNOS |           |                      |                                                |                |                   |
|-------|------------------|-----------|----------------------|------------------------------------------------|----------------|-------------------|
|       | Fator            | Unidade   | Frequência de coleta | Fonte de Informação                            | Área envolvida | Pessoas           |
| 1.1   | Clima            |           |                      |                                                |                |                   |
| 1.1.1 | Temperatura      |           |                      |                                                |                |                   |
|       | Máxima           | °C        | Semanal              | Instituto de meteorologia                      | Planejamento   | a serem definidas |
|       | Média            | °C        | Semanal              | Instituto de meteorologia                      | Planejamento   | a serem definidas |
|       | Mínima           | °C        | Semanal              | Instituto de meteorologia                      | Planejamento   | a serem definidas |
| 1.1.2 | Dia de sol       | SIM / NÃO | Diário               | Instituto de meteorologia / Pessoas envolvidas | Planejamento   | a serem definidas |

| 2   | FATORES INTERNOS         |           |                      |                                |                          |                   |
|-----|--------------------------|-----------|----------------------|--------------------------------|--------------------------|-------------------|
|     | Fator                    | Unidade   | Frequência de coleta | Fonte de Informação            | Área envolvida           | Pessoas           |
| 2.1 | Stock Out                | SIM / NÃO | Semanal              | Sistemas e pessoas das filiais | Planejamento             | a serem definidas |
| 2.2 | Ações de marketing       | SIM / NÃO | Semanal              | Áreas envolvidas               | Marketing                | a serem definidas |
| 2.3 | Ações de trade marketing | SIM / NÃO | Semanal              | Áreas envolvidas               | Trade Marketing e Vendas | a serem definidas |

## 9.2 Análises futuras

Definindo os fatores que seriam importantes para uma análise mais detalhada do comportamento das vendas da Kibon e estruturando o procedimento para a coleta de dados, outros estudos, além daqueles realizados neste trabalho poderiam ser desenvolvidos.

Para comprovar a existência de uma relação significativa entre os fatores, cujo histórico de ocorrências espera-se construir, pode ser feita uma análise de correlação, avaliando o coeficiente, o  $R^2$  e o valor  $P$ , conforme realizado no capítulo de análise.

Vale ressaltar que muitos dos fatores apresentados no capítulo anterior tem como unidade não um valor numérico, mas o atributo “SIM” ou “NÃO”.

Neste caso, uma análise utilizando as chamadas variáveis binárias pode ser realizada.

De acordo com HOFFMANN (1977), variáveis binárias são aquelas que assumem somente dois valores distintos, geralmente zero ou um. Elas seriam usadas para representar variáveis qualitativas, por exemplo, e a equação de regressão seria dada por:

$$Y_i = a + cZ_i + bX_i + e_i \quad (31)$$

onde  $Z_j$  é a variável binária.

Considerando que  $Z_j$  assume valor um ( $Z_j = 1$ ) no primeiro período e valor zero ( $Z_j = 0$ ) no segundo período, a equação que representaria o modelo seria:

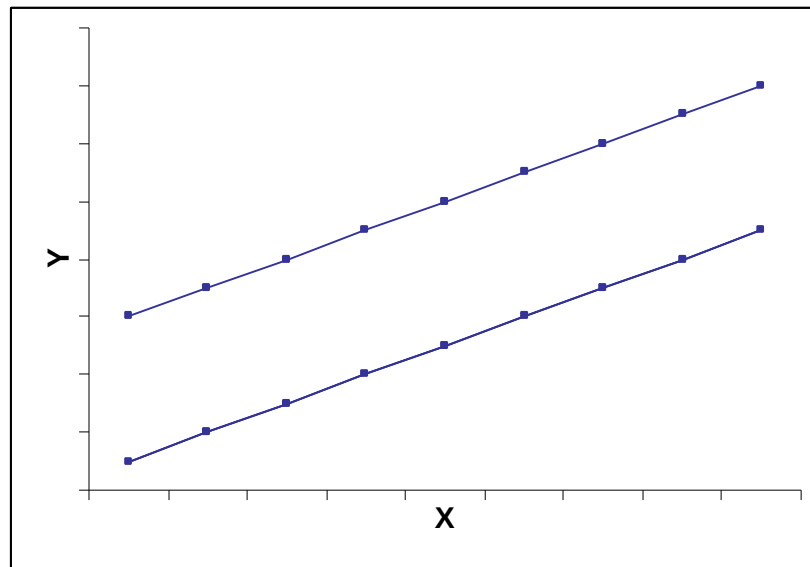
$$Y = (a + c) + bX + e \quad (32)$$

para o período um e:

$$Y = a + bX + e \quad (33)$$

para o segundo período.

Graficamente, a variável binária altera o nível da reta, gerando um par de retas paralelas que representariam a série.



*Figura 9.2 - Representação de uma série qualquer com variável binária*

Dessa maneira, uma variável binária poderia ser atribuída ao fator “dias de sol”, *stock out* e a cada uma das ações de marketing e trade marketing. Por exemplo: nos períodos em que uma propaganda estivesse sendo veiculada na televisão ela assumiria o valor um; nos demais períodos, o valor dessa variável seria zero.

Um modelo de regressão linear que considerasse como variável todas essas citadas anteriormente, mais a temperatura e a própria venda das semanas anteriores, uma vez que foi comprovada a existência dessa relação, poderia representar melhor o comportamento de vendas da Kibon. Conseqüentemente, a equação encontrada pode gerar valores de previsão mais próximos ao valor real.

## Conclusão



## 10. CONCLUSÃO

Para a realização deste trabalho, a autora encontrou algumas dificuldades com relação ao levantamento dos dados utilizados, uma vez que foram levantados dados históricos de 2002 até 2004.

Além disso, conforme já enfatizado anteriormente, o estudo completo do comportamento de vendas da Kibon foi prejudicado devido a falta de dados a respeito de ações e fatores que acredita-se que sejam impactantes para as vendas da empresa.

Outra dificuldade encontrada, citada também por KIYUZATO (2001), foi a falta de comunicação entre as áreas da empresa envolvidas no processo de planejamento da demanda, o que aumentou a dificuldade de coleta de dados.

No entanto, mesmo diante desses problemas, acredita-se que foi possível a realização de um estudo com bastante profundidade estatística em relação aos dados que estavam disponíveis.

O estudo confirmou a hipótese levantada pela empresa da existência de influência da temperatura sobre as vendas, mas mostrou que deve-se ter cuidado ao se afirmar tal constatação, uma vez que a relação só é realmente significativa para algumas regiões do país e mais forte para algumas famílias de produtos.

A análise de autocorrelação realizada evidenciou o fato de que existe também uma forte influência das vendas de períodos anteriores sobre as vendas atuais da empresa.

Diante das dificuldades encontradas e dos resultados obtidos, pode-se ainda levantar outras oportunidades de estudo das vendas da empresa, enfatizando a necessidade de uma coleta de dados mais precisa por parte da empresa.

Sendo assim, acredita-se que os objetivos deste trabalho foram cumpridos, uma vez que proporcionou à empresa um conhecimento mais profundo, baseado em métodos estatísticos, de suas vendas.

Este trabalho também contribuiu para o desenvolvimento acadêmico e pessoal da autora, uma vez que proporcionou a realização do que pode ser considerado um projeto em uma



empresa real, necessitando um conhecimento da organização e da forma de trabalho na mesma.

Por fim, através deste estudo, a autora pode aplicar numa situação real diversos conceitos e métodos aprendidos durante o curso de Graduação em Engenharia de Produção, o que comprova mais uma vez a importância deste trabalho em sua formação pessoal, acadêmica e na construção de seu futuro profissional.

**Anexos**



## 11. ANEXOS

### ANEXO A - Medidas de erro de previsão

Para medir a qualidade das previsões, ou seja, quantificar a acuracidade das mesmas, diversos indicadores podem ser utilizados, tais como o erro médio, o erro absoluto médio, o erro quadrático médio, entre outros que serão apresentados nesse anexo A.

Antes de apresentar as fórmulas que representam os erros citados anteriormente, é preciso definir o chamado erro de previsão, que seria a medida mais simples para medir a qualidade de uma previsão de demanda. MAKRIDAKIS (1998) define este erro como sendo:

$$e_t = Y_t - F_t \quad (34)$$

onde  $e_t$  é o chamado erro de previsão, dado pela diferença entre o valor real observado ( $Y_t$ ) e o previsto ( $F_t$ ) no período  $t$ .

Dessa maneira, o erro médio é dado por:

$$EM = \frac{1}{n} \sum_{i=1}^n e_t \quad (35)$$

De acordo com MAKRIDAKIS (1998), o erro médio indica apenas se o erro de previsão está sendo “para cima” ou “para baixo”, o chamado *forecast bias*, mas não é adequado para dizer a magnitude desses erros. Isso porque os erros positivos e negativos tendem a anular uns aos outros.

Para se conhecer então essa magnitude dos erros, o uso do Erro Absoluto Médio ou do Erro Quadrático Médio é mais adequado.

$$EAM = \frac{1}{n} \sum_{i=1}^n |e_t| \quad (36)$$

$$EQM = \frac{1}{n} \sum_{i=1}^n e_i^2 \quad (37)$$

Ao invés de se calcular os erros sobre o valor de erro médio, pode-se optar por utilizar o erro percentual dado por:

$$EP_t = \left( \frac{Y_t - F_t}{Y_t} \right) \times 100 \quad (38)$$

Dessa maneira, pode-se calcula então o Erro Percentual Médio e o Erro Absoluto Percentual Médio:

$$EPM = \frac{1}{n} \sum_{i=1}^n EP_t \quad (39)$$

$$EAPM = \frac{1}{n} \sum_{i=1}^n |EP_t| \quad (40)$$

MAKRIDAKIS (1998) ressalta o fato de que a melhor maneira de se medir a qualidade de um modelo de previsão não é simplesmente calculando os erros apresentados anteriormente. O ideal é comparar as medidas de erro (seja ela o erro médio, o erro percentual ou o erro percentual médio) entre os próprios modelos de previsão.

Para tanto, sugere aplicar um chamado "modelo ingênuo", que consiste em simplesmente definir como sendo a precisão do período seguinte o valor real mais recente, e comparar a medida de erro entre este modelo e o modelo escolhido.

Um método eficaz para fazer tal comparação seria o da **Estatística-U de Theil**. Além de ser adequado para comparar os resultados de um modelo ingênuo de previsão e outro modelo escolhido, MAKRIDAKIS (1998) apresenta o valor *U* de Theil como uma medida de erro que dá maior peso aos maiores erros, como no caso do erro quadrático médio.

$$U = \sqrt{\frac{\sum_{t=1}^{n-1} \left( \frac{F_{t+1} - Y_{t+1}}{Y_t} \right)^2}{\sum_{t=1}^{n-1} \left( \frac{Y_{t+1} - Y_y}{Y_t} \right)^2}} \quad (41)$$

Os valores de  $U$  encontrados através da fórmula (41) indicam a qualidade do método de previsão utilizado. Segundo MAKRIDAKIS (1998), se:

$U = 1$ : o método ingênuo é tão eficiente quanto o método utilizado.

$U < 1$ : o método utilizado é melhor do que o método ingênuo. Quanto menor o valor de  $U$ , melhor o método é em comparação com o método ingênuo.

$U > 1$ : o método ingênuo apresenta melhores resultado do que o método utilizado.

## ANEXO B – Modelo de Suavização Exponencial

Em muitos casos de previsão de demanda a partir de dados históricos, não é desejável atribuir o mesmo peso a um dado mais antigo e a um mais recente, como feito nos métodos de médias, uma vez que se parte do pressuposto de que os dados mais recentes são mais confiáveis por refletir melhor o padrão de demanda presente.

Nestes casos, pode-se utilizar o chamado método de suavização exponencial; nele, aos dados mais recentes são atribuídos pesos maiores, que vão decaindo exponencialmente, o que explica o nome deste método.

Para tanto, é necessário definir uma série de parâmetros que serão responsáveis por atribuir esses diferentes pesos para os dados, e existem basicamente três métodos de suavização exponencial, dependendo do número de parâmetros utilizados: suavização simples, método de *Holt* e método de *Holt-Winters*.

Além do fato de atribuir pesos maiores para dados mais recentes, HAX (1984) apresenta também como sendo vantagens do método de suavização exponencial os seguintes aspectos, que explicariam a ampla utilização do mesmo:

1. Relativa simplicidade do método
2. Razoável acuracidade
3. Eficiência do ponto de vista computacional

### *Suavização Exponencial Simples*

Neste primeiro método de suavização exponencial, a previsão do período anterior e o erro desta previsão relativo ao valor real observado são utilizados para se determinar a previsão do período em questão:

$$\boxed{F_{t+1} = F_t + a(Y_t - F_t)} \quad (42)$$

onde  $a$  é a chamada constante de suavização e varia entre 0 e 1.

De acordo com SANTORO (2003), valores de  $a$  próximos de 0 indicam que a demanda é mais afetada por variações aleatórias e pouco ajuste sobre o valor da previsão do período anterior é realizado.

Já nos casos em que o valor de  $a$  é mais próximo de 1, admite-se uma forte tendência na demanda, e um grande ajuste sobre a previsão do período  $t-1$  é realizado.

Outra maneira de se escrever a equação anterior é:

$$\boxed{F_{t+1} = aY_t + (1-a)F_t} \quad (43)$$

Sabendo-se que, por sua vez,

$$\boxed{F_t = aY_{t-1} + (1-a)F_{t-1}} \quad (44)$$

e substituindo a equação (44) em (43) e assim por diante para  $F_{t-1}$ ,  $F_{t-2}$ , até  $F_1$ , teremos:

$$\boxed{F_{t+1} = aY_t + a(1-a)Y_{t-1} + a(1-a)^2Y_{t-2} + a(1-a)^3Y_{t-3} + \dots + a(1-a)^{t-1}Y_1 + (1-a)^tF_1} \quad (45)$$

Observando a equação (45), nota-se que são atribuídos pesos diferentes a cada valor histórico observado: pesos menores para valores mais antigos e pesos maiores para valores mais recentes, de acordo com o princípio básico dos métodos de suavização exponencial.

Observe também na mesma equação que é necessário saber o valor de  $F_1$  para se iniciar o processo de previsão. Este valor de inicialização pode ser definido como sendo igual ao valor da primeira observação real da série.

$$\boxed{F_1 = Y_1} \quad (46)$$

MAKRIDAKIS (1998) ainda sugere que o valor inicial de  $F_1$  pode ser definido como sendo a média dos quatro ou cinco primeiros valores da série.

#### *Método Linear de Holt*

Neste método, são utilizadas duas constantes,  $a$  e  $\beta$ , para se determinar a demanda do período em questão. A adição de uma constante permite prever séries com a presença de uma tendência, através das equações (47), (48) e (49).

$$\boxed{L_t = aY_t + (1-a)(L_{t-1} + b_{t-1})} \quad (47)$$

$$\boxed{b_t = b(L_t - L_{t-1}) + (1 - b)b_{t-1}} \quad (48)$$

$$\boxed{F_{t+m} = L_t + b_t m} \quad (49)$$

onde  $L_t$  representa o valor base da série no período  $t$ ,  
 $b_t$  representa o valor de tendência no período  $t$ ,  
 $\beta$  é uma constante que varia entre 0 e 1 e  
 $m$  é o número de períodos futuros que se deseja prever.

Para o caso da aplicação deste método, é necessário determinar os valores de inicialização  $L_1$  e  $b_1$ :

$$\boxed{L_1 = Y_1} \quad (50)$$

$$\boxed{b_1 = Y_2 - Y_1} \quad (51) \text{ ou}$$

$$\boxed{b_1 = (Y_4 - Y_1)/3} \quad (52)$$

### *Método de Holt-Winter*

O método de *Holt-Winter* é uma extensão do método explicado no item anterior, adicionando uma terceira equação e também uma constante para a sazonalidade.

Dessa maneira, a demanda de um período pode ser prevista através das equações (53), (54), (55) e (56).

$$\boxed{L_t = a \frac{Y_t}{S_{t-s}} + (1 - a)(L_{t-1} + b_{t-1})} \quad (53)$$

$$\boxed{b_t = b(L_t - L_{t-1}) + (1 - b)b_{t-1}} \quad (54)$$

$$\boxed{S_t = g \frac{Y_t}{L_t} + (1 - g)S_{t-s}} \quad (55)$$

$$\boxed{F_{t+m} = (L_t + b_t m)S_{t-s+m}} \quad (56)$$



onde  $S_t$  representa a sazonalidade do período  $t$ ,  
 $\beta$  é uma constante e varia entre 0 e 1 e  
 $s$  é o número de períodos que compõem um ciclo completo de sazonalidade.

Como nos casos dos métodos anteriores, é preciso determinar os valores de inicialização  $L_t$ ,  $b_t$  e  $S_t$ . Para tanto, MAKRIDAKIS (1998) apresenta as seguintes equações:

$$L_s = \frac{1}{s} \sum_{i=1}^s Y_i \quad (57)$$

$$b_s = \frac{1}{s} \sum_{i=1}^s \frac{Y_{s+i} - Y_i}{s} \quad (58)$$

$$S_t = \frac{Y_t}{L_s} \quad (59)$$

Em todos os métodos de suavização exponencial apresentados anteriormente, os parâmetros  $\alpha$ ,  $\beta$  e  $\gamma$  devem ser definidos de maneira a minimizar as medidas de erro, seja ela o de erro absoluto percentual ou de erro quadrático, o que pode ser feito através de um algoritmo de otimização não-linear.

MAKRIDAKIS (1998) sugere ainda que diversos valores para as constantes sejam testados, para se selecionar aqueles que apresentarem o menor erro de previsão.

## **Referência Bibliográfica**



## 12. REFERÊNCIA BIBLIOGRÁFICA

ACNIELSEN, **Ruptura: causas e impactos na cadeia de abastecimento e no comportamento do consumidor**, São Paulo, 2004.

BALLOU, RONALD H., **Gerenciamento da Cadeia de Suprimentos**, 4a edição, Porto Alegre: Bookman, 2001.

DIEBOLD, FRANCIS X., **Elements of forecasting**, Cincinnati: South-Western College Publishing, 1998.

HAX, ARNOLDO C. & CANDEA, DAN, **Production and Inventory Management**, Prentice-Hall Inc., 1984.

HOFFMANN, RODOLFO & VIEIRA, SÔNIA, **Análise de Regressão**, São Paulo: Editora Hucitec, 1977.

KACHIGAN, SAM W., **Multivariate Statistical Analysis**, 2a edição, New York: Radius Press, 1986.

KIYUZATO, LEANDRO, **Aplicação de Modelos de Previsão em empresas de vendas sazonais**, São Paulo, 2001.

MAKRIDAKIS, SPYROS G., **Forecasting: Methods and Applications**, 3a edição, John Wiley & Sons, 1998.

OLIVEIRA, DJALMA DE PINHO REBOUÇAS DE, **Sistemas, organizações e métodos**, 5a edição, São Paulo: Editora Atlas, 1994.

RAMOS, ALBERTO W., **Uma contribuição aos estudos de capacidade de máquina**, Dissertação (Doutorado) – Escola Politécnica da Universidade de São Paulo, São Paulo, 1999.

SANTORO, MIGUEL C., **Planejamento, Programação e Controle da Produção**, Notas de aula, São Paulo, 2003.

WHEELWRIGHT, STEVEN C. & MAKRIDAKIS, SPYROS G., **Forecasting methods for management**, 2a edição, John Wiley & Sons, 1997.