



Universidade Federal De Uberlândia

Faculdade De Matemática

Curso De Estatística

**Marco Aurélio Maia Silva**

**DISTRIBUIÇÃO ESPACIAL E TEMPORAL  
DE VENDAS DE PEQUENOS E MÉDIOS  
VAREJISTAS DO ESTADO DA BAHIA**

UBERLÂNDIA – MINAS GERAIS  
2015



Universidade Federal De Uberlândia

Faculdade De Matemática

Curso De Estatística

# **DISTRIBUIÇÃO ESPACIAL E TEMPORAL DE VENDAS DE PEQUENOS E MÉDIOS VAREJISTAS DO ESTADO DA BAHIA**

*Monografia apresentada ao Programa de Graduação  
em Estatística da Universidade Federal de  
Uberlândia, como requisito essencial para obtenção  
do Grau de Bacharel em Estatística.*

Orientador: Professor Ednaldo Carvalho Guimarães

UBERLÂNDIA – MINAS GERAIS  
2015

Universidade Federal De Uberlândia

Faculdade De Matemática

Curso De Estatística

**DISTRIBUIÇÃO ESPACIAL E TEMPORAL  
DE VENDAS DE PEQUENOS E MÉDIOS  
VAREJISTAS DO ESTADO DA BAHIA**

Monografia apresentada à Universidade Federal de Uberlândia,  
perante a banca de examinadores abaixo, como requisito essencial  
para obtenção do Grau de Bacharel em Estatística. Aprovada dia  
07/12/2015.

Banca Examinadora:

Professor Ednaldo Carvalho Guimarães

Professor Marcelo Tavares

Maria Imaculada de Sousa Silva

Tiago Moreira Vargas

# Resumo

A estatística é uma ciência muito útil na interpretação dos dados e auxílio na tomada de decisões. A geoestatística, por exemplo, é uma ferramenta utilizada para estudar a variabilidade espacial. Ela permite a interpretação dos resultados com base na estrutura da variabilidade natural dos atributos avaliados, e pondera a dependência espacial dentro do intervalo de amostragem, já a análise de séries temporais modela o comportamento temporal da variável e utiliza do modelo para realizar projeções. Assim, o objetivo deste trabalho foi utilizar a análise geoestatística para fazer um mapeamento das vendas médias com base nos últimos 12 meses de 171 clientes da Bahia e realizar uma projeção das vendas até 10/2016 com base nas informações dos últimos 5 anos. Para isto, os bancos de dados utilizados neste trabalho foram fornecidos por uma empresa atacadista de Uberlândia. O procedimento para análise foi realizado em três fases de análise: descritiva, geoestatística e temporal. Os softwares utilizados para as análises foram o geoR, SPSS, SAS Visual Analytics e GS +. Na análise geoestatística verificou-se que para a variável analisada não houve dependência espacial. Na análise temporal foram ajustados vários modelos para explicar a série e os modelos com melhor ajuste foram SARIMA (0,1,1) (1,1,0) e Holt-Winters aditivo, sendo este último utilizado para as projeções. A exploração estatística de dados das empresas se constituem ferramentas de grande importância no planejamento e tomada de decisões das mesmas.

Palavras-chave: estatística espacial; séries temporais, interpolação; *forecast*.

# Abstract

Statistics is a very useful science in the interpretation of data and aid in decision making. Geostatistics, for example, is a tool used to study the spatial variability. It allows the interpretation of the results based on the structure of the natural variability of the evaluated attributes, and considers the spatial dependence within the sampling interval, since the time series analysis models the temporal behavior of the variable and the model used to make projections. The objective of this work was to use the geostatistic analysis to make a mapping of the average sales based on the past 12 months 171 customers of Bahia and hold a sales projection to 10/2016 based on the information of the last five years. For this, the databases used in this study were provided by a wholesale company of Uberlândia. The procedure for analysis was conducted in three phases of analysis: descriptive, geostatistics and time series. The software used for analyzes were geoR, SPSS, SAS Visual Analytics and GS +. In geostatistic analysis it was found that for the analyzed variable was no spatial dependence. In the temporal analysis were adjusted several models to explain the series and the best fit models were SARIMA ( 0,1,1 ) ( 1,1,0 ) and Holt -Winters additive, the latter being used for the projections. Statistical operating data of the companies are very important tools in planning and making decisions them.

Key-words: spatial statistics; time series; interpolation; forecast.

# Sumário

|                                      |    |
|--------------------------------------|----|
| Resumo .....                         | 3  |
| Abstract.....                        | 5  |
| Lista de Figuras .....               | 7  |
| Lista de Tabelas .....               | 9  |
| Introdução e Justificativa.....      | 10 |
| 1.1    Objetivo Geral.....           | 12 |
| 1.2    Objetivos específicos .....   | 13 |
| Referencial Teórico .....            | 14 |
| 2.1    Análises Geoestatísticas..... | 14 |
| 2.2    Análises Temporais.....       | 16 |
| Materiais e Métodos .....            | 19 |
| 3.1    Estatística Descritiva.....   | 19 |
| 3.2    Geoestatística .....          | 19 |
| 3.3    Estatística Temporal .....    | 20 |
| Resultados e Discussão.....          | 21 |
| 4.1    Dados geoestatísticos .....   | 21 |
| 4.2    Dados temporais .....         | 25 |
| Conclusão .....                      | 32 |
| Referências .....                    | 33 |

# Lista de Figuras

|                                                                                                                                                                            |    |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| FIGURA 1: BOXPLOT DAS VENDAS MÉDIAS DE 171 CLIENTES VAREJISTAS DE UM ATACADISTA DE UBERLÂNDIA-MG, NO ESTADO DA BAHIA, NO PERÍODO DE 09/14 A 09/15.....                     | 22 |
| FIGURA 2: HISTOGRAMA DAS VENDAS MÉDIAS DE 171 CLIENTES VAREJISTAS DE UM ATACADISTA DE UBERLÂNDIA-MG, NO ESTADO DA BAHIA, NO PERÍODO DE 09/14 A 09/15.....                  | 22 |
| FIGURA 3: GRÁFICO DESCRITIVO ESPACIAL DAS VENDAS MÉDIAS DE 171 CLIENTES VAREJISTAS DE UM ATACADISTA DE UBERLÂNDIA-MG, NO ESTADO DA BAHIA, NO PERÍODO DE 09/14 A 09/15..... | 23 |
| FIGURA 4: SEMIVARIOGRAMA DAS VENDAS MÉDIAS DE 171 CLIENTES VAREJISTAS DE UM ATACADISTA DE UBERLÂNDIA-MG, NO ESTADO DA BAHIA, NO PERÍODO DE 09/14 A 09/15.....              | 23 |
| FIGURA 5: LOCALIZAÇÃO GEOGRÁFICA DOS CLIENTES VAREJISTAS DE UM ATACADISTA DE UBERLÂNDIA-MG, NO ESTADO DA BAHIA, NO PERÍODO DE 09/14 A 09/15.....                           | 24 |
| FIGURA 6: MAPA POR INTERPOLAÇÃO IDW DOS CLIENTES VAREJISTAS DE UM ATACADISTA DE UBERLÂNDIA-MG, NO ESTADO DA BAHIA, NO PERÍODO DE 09/14 A 09/15.....                        | 25 |
| FIGURA 7: BOXPLOT DAS VENDAS TOTAIS MENS AIS DE CLIENTES VAREJISTAS DE UM ATACADISTA DE UBERLÂNDIA-MG, NO ESTADO DA BAHIA, NO PERÍODO DE 01/10 A 09/15.....                | 26 |
| FIGURA 8: HISTOGRAMA DAS VENDAS TOTAIS MENS AIS DE CLIENTES VAREJISTAS DE UM ATACADISTA DE UBERLÂNDIA-MG, NO ESTADO DA BAHIA, NO PERÍODO DE 01/10 A 09/15.....             | 26 |

|                                                                                                                                                                                                                                                             |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| FIGURA 9: COMPORTAMENTO DA SÉRIE TEMPORAL DAS VENDAS TOTAIS MENSAIS DE<br>CLIENTES VAREJISTAS DE UM ATACADISTA DE UBERLÂNDIA-MG, NO ESTADO DA<br>BAHIA, NO PERÍODO DE 01/10 A 09/15.....                                                                    | 27 |
| FIGURA 10: COMPARATIVO DOS MODELOS AJUSTADOS E VALORES OBSERVADOS DAS<br>VENDAS TOTAIS MENSAIS DE CLIENTES VAREJISTAS DE UM ATACADISTA DE<br>UBERLÂNDIA-MG, NO ESTADO DA BAHIA, NO PERÍODO DE 01/10 A 09/15. ....                                           | 29 |
| FIGURA 11: AUTOCORRELAÇÃO E AUTOCORRELAÇÃO PARCIAL DOS RESÍDUOS PARA O<br>MODELO HOLT-WINTERS DAS VENDAS TOTAIS MENSAIS DE CLIENTES VAREJISTAS DE<br>UM ATACADISTA DE UBERLÂNDIA-MG, NO ESTADO DA BAHIA, NO PERÍODO DE 01/10<br>A 09/15. ....               | 30 |
| FIGURA 12: SÉRIE OBSERVADA, MODELO AJUSTADO, <i>FORECAST</i> E LIMITES DO INTERVALO<br>DE CONFIANÇA DE 95%, DAS VENDAS TOTAIS MENSAIS DE CLIENTES VAREJISTAS DE<br>UM ATACADISTA DE UBERLÂNDIA-MG, NO ESTADO DA BAHIA, NO PERÍODO DE 01/10<br>A 09/15. .... | 30 |

# Lista de Tabelas

|                                                                                                                                                                                                                  |    |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| TABELA 1: ESTATÍSTICA DESCRITIVA DAS VENDAS MÉDIAS DE 171 CLIENTES VAREJISTAS DE UM ATACADISTA DE UBERLÂNDIA-MG, NO ESTADO DA BAHIA, NO PERÍODO DE 09/14 A 09/15.....                                            | 21 |
| TABELA 2: ESTATÍSTICA DESCRITIVA DOS DADOS HISTÓRICOS DAS VENDAS TOTAIS MENSAIS DE CLIENTES VAREJISTAS DE UM ATACADISTA DE UBERLÂNDIA-MG, NO ESTADO DA BAHIA, NO PERÍODO DE 01/10 A 09/15. ....                  | 25 |
| TABELA 3: TESTES PARA AVALIAR PRESENÇA DE ESTACIONARIEDADE E TENDÊNCIA DAS VENDAS TOTAIS MENSAIS DE CLIENTES VAREJISTAS DE UM ATACADISTA DE UBERLÂNDIA-MG, NO ESTADO DA BAHIA, NO PERÍODO DE 01/10 A 09/15. .... | 27 |
| TABELA 4: COMPARATIVO ENTRE OS MODELOS DE SÉRIES TEMPORAIS DAS VENDAS TOTAIS MENSAIS DE CLIENTES VAREJISTAS DE UM ATACADISTA DE UBERLÂNDIA-MG, NO ESTADO DA BAHIA, NO PERÍODO DE 01/10 A 09/15. ....             | 28 |
| TABELA 5: PREDIÇÃO DOS MODELOS AJUSTADOS DE SÉRIES TEMPORAIS DAS VENDAS TOTAIS MENSAIS DE CLIENTES VAREJISTAS DE UM ATACADISTA DE UBERLÂNDIA-MG, NO ESTADO DA BAHIA, NO PERÍODO DE 01/10 A 09/15.....            | 28 |

## Introdução e Justificativa

Varejo é a palavra utilizada para indicar os setores de comércio que vendem diretamente aos consumidores finais, sem intermediários. Assim, para Parente (2000), o varejo consiste em processos tanto de vendas de produtos quanto de serviços que atendem uma necessidade e a satisfação pessoal do consumidor.

Desta forma, conhecer o consumidor e saber identificar suas necessidades e interesses de consumo é fundamental para qualquer empresa, já que para vender, as mesmas (empresas) devem ter a finalidade de satisfazer as necessidades dos consumidores (SOLOMON, 2007).

Assim, os profissionais das empresas procuram aprender caminhos que ajudam a influenciar na venda para os consumidores, e esses caminhos compreendem ações como: ajudar a identificar as necessidades e exigências dos consumidores; fornecer para a empresa as opiniões dos consumidores acerca dos produtos fornecidos e possibilidades de mudanças e melhorias; e por fim, acompanhar na pós-venda o nível de satisfação dos consumidores (KOTLER, 1998).

A estatística é uma ferramenta matemática usada largamente em empresas, principalmente quando se fala em previsões, obtenção de conclusões válidas e em tomada de decisões razoáveis baseadas em análise e interpretação de dados (MATAR et al, 1999; COBRA, 2009), entretanto este ramo matemático poderia ser bem mais explorado pelas empresas.

A geoestatística, por exemplo, é uma ferramenta estatística utilizada para estudar a variabilidade espacial. Ela permite a interpretação dos resultados com base na estrutura da variabilidade natural dos atributos avaliados, e pondera a dependência espacial dentro do intervalo de amostragem (DE OLIVEIRA MACHADO et al, 2007).

A função semivariograma é muito utilizada na geoestatística para determinar e explorar os padrões ou continuidades espaciais (SOARES, 2000). Desta forma, o semivariograma é um gráfico que relaciona a semivariância de uma variável qualquer com uma distância ( $h$ ). Os semivariogramas apresentam três importantes parâmetros: (1) o efeito pepita ( $C_0$ ), que se refere ao valor da semivariância para a distância zero e representa o componente da variação ao acaso; (2) o patamar ( $C + C_0$ ), que é o valor da semivariância em que a curva estabiliza sobre um valor constante (é o máximo da semivariância) (DAS, 2003). O patamar é atingido quando a variância dos dados se torna constante com as distâncias entre as amostras e esse parâmetro permite a determinação da distância limite entre a dependência e a independência entre as amostras; (3) o alcance ( $a$ ), que é a distância da origem até onde o patamar atinge valores estáveis, é considerado o limite da dependência espacial da grandeza medida (SILVA, REICHERT, STORCK, FEIJO, 2003).

Após a seleção do semivariograma da variável em estudo e havendo dependência espacial, podem-se interpolar valores em qualquer posição na área estudada, sem tendência e com variância mínima (VIEIRA et al, 1983).

Além da variabilidade espacial também é interessante, nos estudos das vendas a varejo, o conhecimento da variabilidade temporal e, nesse contexto, destaca-se a análise de séries temporais.

As técnicas de análise de séries temporais vêm sendo utilizadas com frequência em várias áreas de pesquisa, como: ciências sociais e políticas, economia, sociologia, história, psicologia, meio ambiente, entre outras, nas quais as observações estão geralmente associadas ao tempo e seus métodos evidenciam características dinâmicas dos fenômenos (SAFADI, 2003).

Uma série histórica pode ser composta por quatro componentes: tendência, sazonalidade, ciclicidade e a variação aleatória denominada de ruído branco.

A tendência de uma série indica o seu comportamento “de longo prazo”, ou seja, se ela sobe, desce, ou permanece estável. Para analisar o efeito da tendência pode ser utilizado o teste de Cox-Stuart, conhecido como teste da tendência ou do sinal (MORETTIN e TOLOI, 1985).

De acordo com Barros (2009), define-se como um fenômeno sazonal aquele que ocorre regularmente em períodos fixos de tempo. Para verificar a existência de sazonalidade na série, pode ser utilizada a função de autocorrelação.

Existem alguns métodos de estimação de séries:

- i. Médias móveis (consiste em calcular a média aritmética das  $k$  observações mais recentes, e a cada período a observação mais antiga é substituída pela mais recente).
- ii. Alisamento exponencial (a ideia é parecida com as médias móveis, porém os pesos das observações decrescem à medida que as observações estão mais “longes do passado”).
- iii. Box-Jenkins (costumam apresentar resultados melhores que os métodos de amortecimento quando a série é relativamente longa e sem muitas oscilações).
- iv. Holt-Winters (descrevem apropriadamente dados em que se verifica a ocorrência de tendência linear, além de componente de sazonalidade (PELEGRINI e FLOGIATTO, 2000), sendo sua aplicação válida pra séries não estacionárias).

Ciclo ou componente cíclica é um movimento oscilatório de longa duração (mais de um ano) que exprime a influência de fatores aleatórios de ação reiterada. Tal componente indica as fases de expansão e contração das atividades econômicas, sendo de duração não fixa. Como este trabalho não possui uma série muito longa, este componente não será avaliado.

A quarta componente da série, chamada de componente de erro, apresenta movimentos ascendentes e descendentes da série após a ocorrência de um efeito de tendência, um efeito cíclico, ou de um efeito sazonal. Nas componentes de erro aparecem flutuações de período curto, com deslocamento inexplicável e geralmente são causadas, entre outros motivos, por eventos políticos e oscilações climáticas imprevisíveis.

## 1.1 Objetivo Geral

O objetivo geral deste trabalho foi fazer um mapeamento das vendas médias dos últimos 12 meses de 171 clientes da Bahia e uma projeção das vendas de todos os clientes da empresa atacadista dos últimos 5 anos.

## 1.2 Objetivos específicos

- Descrever as características das duas bases de dados por meio de análise descritiva.
- Verificar a dependência espacial da concentração de venda dos pequenos e médios varejistas do estado da Bahia.
- Ajustar o modelo de dependência espacial que descreva o comportamento das vendas médias dos clientes do Estado da Bahia por meio de semivariograma.
- Construir mapas que descrevam o comportamento das vendas por meio de interpoladores geoestatísticos (krigagem) ou por meio do IDW (Ponderação do Inverso da Distância).
- Utilizar técnicas de séries temporais para verificar possível presença de tendência ou sazonalidade das vendas.
- Ajustar modelos de séries temporais para as vendas e utilizar os modelos para fazer projeções.

# Capítulo 2

## Referencial Teórico

### 2.1 Análises Geoestatísticas

Os procedimentos de análise geoestatística consistem da utilização da estrutura de variabilidade espacial da variável analisada para se fazer estimativas em locais não amostrados (Vieira et al (1983) e Soares (2000)). Ainda segundo esses autores os procedimentos básicos de análise consistem da análise descritiva, do cálculo da semivariância que define a estrutura de dependência espacial, do ajuste de modelos teóricos de semivariogramas e da utilização desses modelos na estimação por meio de krigagem.

A análise exploratória é feita por meio de análises gráficas (boxplot, histogramas, gráficos de dispersão da variável em relação aos eixos coordenados) e por meio do cálculo de estatísticas básicas (media, quartis, desvio padrão) e tem a finalidade de descrever a variável em análise e detectar valores extremos e possíveis tendências nos dados.

A análise da estrutura de dependência espacial é, geralmente, feita por meio da função semivariâncias. As semivariâncias são obtidas, a princípio, pelo estimador de Matheron que é dado pela equação:

$$\gamma(h) = \left( \frac{1}{2N(h)} \right) \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2, \quad (1)$$

em que  $\gamma(h)$  é a semivariância na distância  $h$ ;  $Z(x)$  é o valor assumido pelo atributo na posição  $x$ ;  $Z(x + h)$  é o valor assumido pelo atributo na posição  $x + h$ ;  $h$  é a distância entre os pontos;  $N(h)$  é o número de pontos na distância  $h$ .

Às semivariâncias obtidas pelos estimadores supracitados serão ajustados os modelos teóricos de semivariograma. Os modelos de semivariograma mais utilizados na literatura são:

- Modelo Esférico:

$$\gamma(h) = C_o + Cx \left( 1,5x \left( \frac{h}{a} \right) - 0,5x \left( \frac{h^3}{a^3} \right) \right), 0 \leq h \leq a \quad (2)$$

$$\gamma(h) = C_o + C, h > a$$

- Modelo Exponencial:

$$\gamma(h) = C_o + Cx \left( 1 - \exp \left( - \frac{(3xh)}{a} \right) \right) \quad (3)$$

- Modelo Gaussiano

$$\gamma(h) = C_o + Cx \left( 1 - \exp \left( - \frac{(3x h^2)}{a^2} \right) \right), \quad (4)$$

em que  $C_o$ ,  $C + C_o$ ,  $a$  e  $h$  são, respectivamente, efeito pepita, patamar, alcance da dependência espacial e distância entre observações.

A seleção do modelo pode ser feita por meio do maior coeficiente de determinação ( $R^2$ ) e a menor soma de quadrados dos resíduos, ou ainda pela validação cruzada (Guimarães, 2000). Ribeiro Junior e Diggle (2001) indicam ainda a possibilidade de utilização dos critérios de informação Akaike (AIC) e critério de informação bayesiana (BIC).

O ajuste de modelos às semivariâncias observadas pode ser realizado a critério do observador (método a sentimento), por mínimos quadrados ordinários, mínimos quadrados ponderados, máxima verossimilhança ou máxima verossimilhança restrita, conforme indicam Ribeiro Junior e Diggle (2001).

Definido o modelo do semivariograma, passa-se para a fase de interpolação por krigagem nos casos em que ocorre a dependência espacial, ou inversa da distância ponderada (IDW) em caso contrário.

A técnica geoestatística denominada krigagem foi definida em VIEIRA et al (1983) e em TRANGMAR et al (1985) como uma técnica de estimação de valores de uma variável aleatória em locais não amostrados, usando as propriedades estruturais do semivariograma e o conjunto de valores experimentais inicial.

O estimador de krigagem determina uma ponderação dos estimadores amostrados, sendo os pesos calculados de forma a minimizar a variância do erro de predição, levando em consideração a configuração espacial da amostra (BRAGA, 1990).

A krigagem é calculada por:

$$Z_x^* = \sum_{i=1}^n \lambda_i Z(x_i), \quad (5)$$

em que  $Z_x^*$  é o valor a ser estimado na posição  $x$ ;  $\lambda_i$  o peso atribuído a observação  $Z(x_i)$ ;  $n$  o número de vizinhos utilizados na estimativa.

A IDW implementa explicitamente o pressuposto de que as coisas mais próximas entre si são mais parecidas do que as mais distantes (JAKOB e YOUNG, 2006).

Para predizer um valor para algum local não medido, o IDW (Ponderação do inverso da distância) usará os valores amostrados à sua volta, que terão um maior peso do que os valores mais distantes, ou seja, cada ponto possui uma influência no novo ponto, que diminui na medida em que a distância aumenta daí seu nome.

O resultado final da interpolação geralmente é expresso em forma de mapa da variável analisada facilitando a interpretação dos resultados.

## 2.2 Análises Temporais

A realização de previsão é um importante fator no processo decisório, tanto no nível estratégico quanto no operacional. Prever níveis de demanda é vital à empresa como um todo, porque fornece as entradas básicas para planejamento e controle de

todas as áreas funcionais, incluindo a logística, o marketing, a produção e as finanças (BALLOU, 2001).

Segundo Slack et. al. (2002) prover a capacitação de satisfazer a demanda atual e futura é uma responsabilidade fundamental da administração da produção. Os autores ainda enfatizam que um equilíbrio adequado entre capacidade e demanda pode gerar altos lucros e clientes satisfeitos, e que o oposto pode ser um desastre.

De acordo com Tubino (2000) as empresas, de uma ou de outra maneira, direcionam suas atividades para o rumo em que acreditam que seu negócio irá se direcionar, sendo assim as técnicas de previsão fundamentais.

Uma série temporal é uma sequência de observações feitas em um período de tempo (BOX, JENKINS e REINZEL, 1994). Se estas observações consecutivas são dependentes uma das outras, é possível conseguir uma previsão (SAMOHYL, ROCHA e MATTOS, 2001), identificando como será o comportamento do evento. Podem ser utilizados modelos como: Holt-Winters Aditivo ou multiplicativos, médias móveis simples, amortecimento exponencial e Box-Jenkins.

Os modelos de Holt-Winters (HW) descrevem apropriadamente dados em que se verifica a ocorrência de tendência linear, além de componente de sazonalidade (PELEGRINI e FLOGIATTO, 2000). Sua aplicação é válida pra séries não-estacionárias.

Morettin e Toloi (2004) comentam que as vantagens desse modelo são: fácil entendimento, aplicação não dispendiosa, adequada para série com padrão de comportamento mais geral. Existem dois modelos:

- Holt-Winters Aditivo: é utilizado quando a amplitude da variação sazonal mantém-se constante, ou seja, a diferença entre o maior e o menor ponto de demanda nos ciclos permanece constante com o passar do tempo.
- Holt-Winters Multiplicativo: é utilizado quando a amplitude da variação sazonal aumenta com o tempo, ou seja, a diferença entre o maior e o menor ponto de demanda nos ciclos cresce com o passar do tempo.

O modelo aditivo é descrito por uma equação principal e três equações de suavização que a compõem. Estas equações serão apresentadas nessa seção juntamente com as equações de previsão, como apresenta Morettin e Toloi (2004).

O modelo aditivo é dado pela seguinte equação:

$$Z_t = \mu_t + T_t + F_t + \varepsilon_t, \quad (6)$$

para  $t = 1, \dots, N$ , onde  $N$  é o tamanho da série, isto é, o número de observações da série.

As equações de suavização são apresentadas a seguir e representam as estimativas do fator sazonal, do nível e da tendência.

$$\begin{aligned}\hat{F}_t &= D(Z_t - \bar{Z}_t) + (1 - D)\hat{F}_{t-s}, \quad 0 < D < 1, \\ \bar{Z}_t &= A(Z_t - \hat{F}_{t-s}) + (1 - A)(\bar{Z}_{t-1} + \hat{T}_{t-1}), \quad 0 < A < 1, \\ \hat{T}_t &= C(\bar{Z}_t - \bar{Z}_{t-1}) + (1 - C)\hat{T}_{t-1}, \quad 0 < C < 1,\end{aligned}\quad (7)$$

onde  $A$ ,  $C$  e  $D$  são as constantes de suavização.

O valor ótimo das constantes de suavização é determinado pelo vetor  $(D, A, C)$  que minimiza a soma dos erros quadráticos de previsão.

As equações de previsão para o modelo aditivo são dadas pelas seguintes expressões as quais dependem do valor da sazonalidade  $s \in \mathbb{N}$ .

$$\hat{Z}_t(h) = \bar{Z}_t + h\hat{T}_t + \hat{F}_{t+h-s}, \quad (8)$$

onde  $h = 1, 2, \dots, s$ .

$$\hat{Z}_t(h) = \bar{Z}_t + h\hat{T}_t + \hat{F}_{t+h-2s}, \quad (9)$$

onde  $h = s+1, \dots, 2s$ .

Os demais métodos são também descritos em Moretin e Toloi (2004) e não foram abordados com detalhes neste estudo em função do mesmo ter focado no modelo de Holt-Winters.

# Capítulo 3

## Materiais e Métodos

Os bancos de dados utilizados neste trabalho foram fornecidos por uma empresa atacadista de Uberlândia. Para a análise geoestatística, foi utilizada a média de vendas dos últimos 12 meses, de 171 clientes da Bahia, com perfil varejista de pequeno e médio porte. Para análise de séries, foi utilizado o valor total de vendas mensal de todos os clientes da empresa atacadista dos últimos 5 anos.

O procedimento para análise foi realizado em três fases: análise descritiva, geoestatística e temporal.

Os softwares utilizados para as análises foram o geoR (CHRISTENSEN, 2002), SPSS (IBM, 2011), SAS Visual Analytics (KEIM, 2010) e GS + (GS +, 2000).

### 3.1 Estatística Descritiva

Foi feito inicialmente uma análise descritiva das duas bases para identificar a dispersão dos dados em relação à curva normal, o coeficiente de variação, as medidas de posição e a simetria tendo a finalidade de explorar e descrever a variável em análise.

### 3.2 Geoestatística

Após o processo de análise descritiva geral, foi realizada a análise descritiva espacial, visando detectar tendências nos dados. Em seguida foram construídos os semivariogramas, realizada a interpolação por krigagem e feito o mapeamento para a venda.

O cálculo das semivariâncias e a escolha do modelo de semivariograma foram feitos através do software R (RIBEIRO e DIGGLE, 2001), que utiliza o pacote geoR.

A interpolação por meio de IDW foi feita no programa GS+ (GS+, 2000).

Toda a análise geoestatística seguiu os procedimentos descritos em Vieira et al (1983), Ribeiro Junior e Diggle (2001) e Soares (2000).

### **3.3 Estatística Temporal**

Após o processo de análise descritiva, o próximo passo foi avaliar se há a presença de tendência (pelo teste de Cox-Stuart) ou sazonalidade nos dados. Após esta avaliação, foram ajustados alguns modelos e os dois que obtiveram melhor ajuste foram utilizados para comparação. Os modelos foram ajustados no SPSS (IBM, 2011).

Os procedimentos da análise de séries temporais seguiram a metodologia descrita em Morettin e Toloi (2004).

## Capítulo 4

# Resultados e Discussão

### 4.1 Dados geoestatísticos

Na análise descritiva dos dados utilizados para análise geoestatística, verificou-se que a variável analisada apresenta valores semelhantes entre a média e a mediana, indicando tendência a simetria da distribuição da variável analisada, como mostrado na Tabela 1. A simetria é desejável para a utilização do estimador clássico de semivariância, conforme argumentam (VIEIRA, 2000 e GUIMARÃES, 2004).

Tabela 1: Estatística descritiva das vendas médias de 171 clientes varejistas de um atacadista de Uberlândia-MG, no estado da Bahia, no período de 09/14 a 09/15.

| Medidas       | Valores |
|---------------|---------|
| Mínimo        | 352     |
| Máximo        | 3.941   |
| Média         | 2.019   |
| Mediana       | 1.852   |
| Desvio Padrão | 962     |
| Curtose       | -0,865  |
| Assimetria    | 0,290   |
| CV            | 47,65%  |

Percebe-se também coeficiente de variação moderado de 47,65%, uma assimetria à direita e uma curtose platicúrtica, ou seja, com caudas mais curtas e menor probabilidade que a normal de ter outliers, como pode ser visto no boxplot (Figura 1).

Segundo Vanni (1998), coeficiente de variação maior que 35% mostra que a série é heterogênea, e a média tem pouco significado. Este fato era esperado por se tratar de uma área de grande extensão territorial e mostra a necessidade do uso de estimativas que considerem a posição espacial das observações, como é o caso da krigagem.

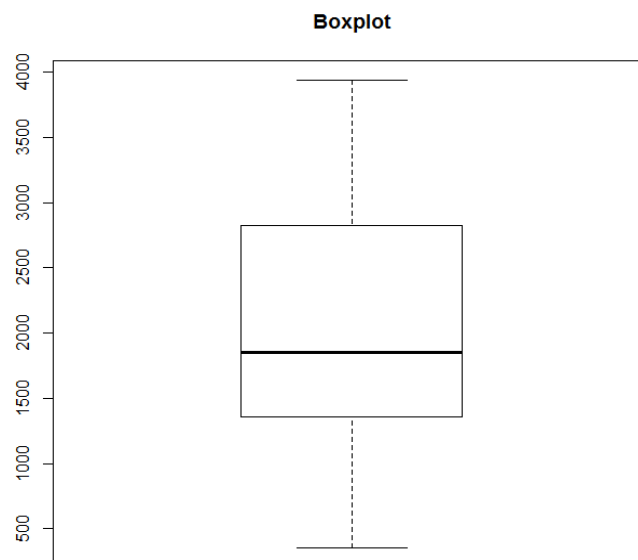


Figura 1: Boxplot das vendas médias de 171 clientes varejistas de um atacadista de Uberlândia-MG, no estado da Bahia, no período de 09/14 a 09/15.

No histograma (Figura 2), nota-se que aproximadamente 58% dos clientes possuem valor de compra de até R\$ 2.000. Isto mostra que mais da metade dos clientes desta região são pequenos, ou seja, possuem um número pequeno de *check-outs*.

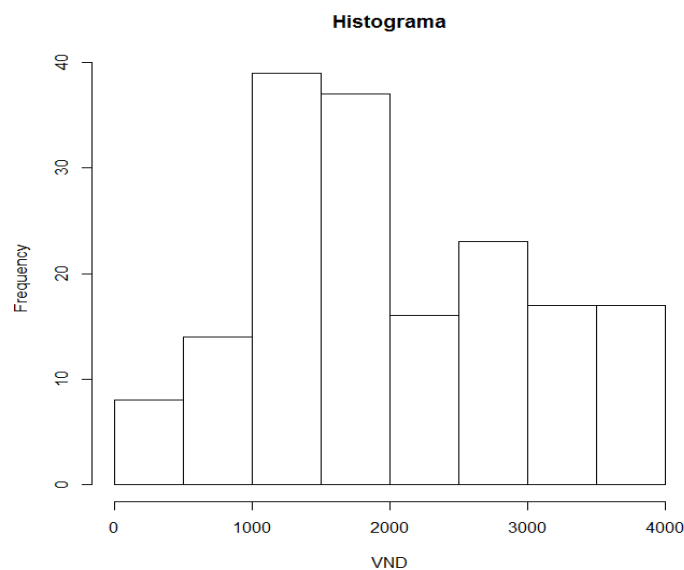


Figura 2: Histograma das vendas médias de 171 clientes varejistas de um atacadista de Uberlândia-MG, no estado da Bahia, no período de 09/14 a 09/15.

Na análise descritiva espacial percebe-se que, apesar dos quartis se distribuírem por toda a área, os clientes com menor faixa de venda se concentram um pouco mais na parte inferior esquerda do mapa (Figura 3).

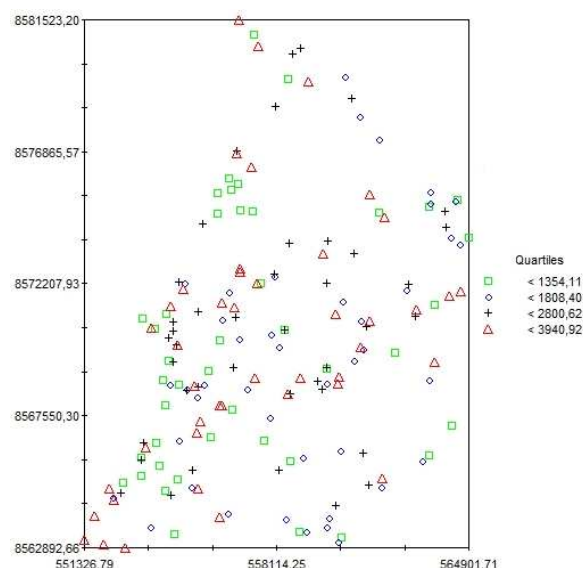


Figura 3: Gráfico descritivo espacial das vendas médias de 171 clientes varejistas de um atacadista de Uberlândia-MG, no estado da Bahia, no período de 09/14 a 09/15.

O semivariograma experimental da variável “vendas”, apresentado na Figura 4, indica comportamento aleatório, ou seja, não ocorre uma estrutura de dependência espacial e desta forma tem-se o efeito pepita puro.

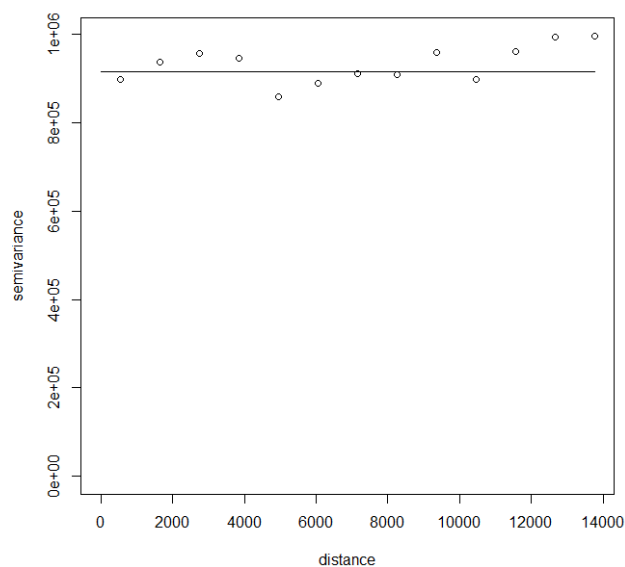


Figura 4: Semivariograma das vendas médias de 171 clientes varejistas de um atacadista de Uberlândia-MG, no estado da Bahia, no período de 09/14 a 09/15.

Este fato acontece devido à independência dos dados, que neste caso pode ter sido causada pela grande distância geográfica entre as observações, como pode ser visualizado na Figura 5. Pela legenda, quanto mais azul o círculo, maior o valor de venda, e quanto mais vermelho, menor.

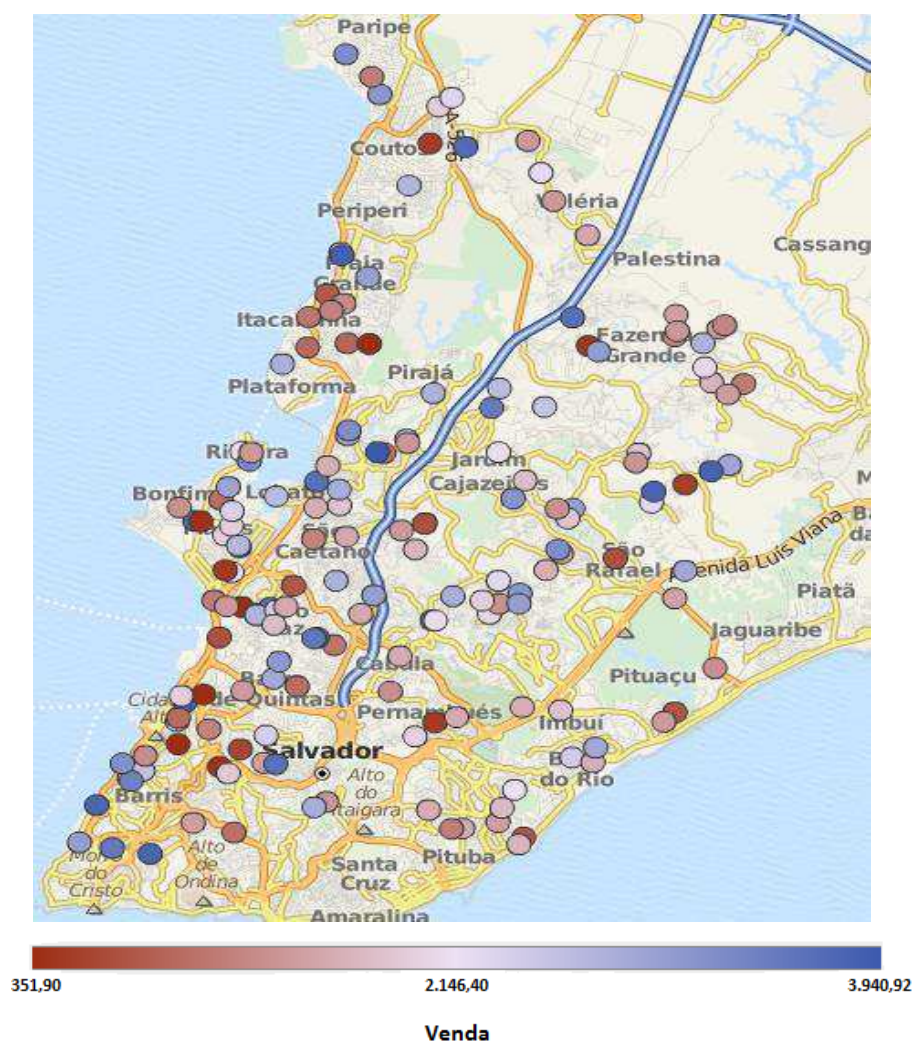


Figura 5: Localização geográfica dos clientes varejistas de um atacadista de Uberlândia-MG, no estado da Bahia, no período de 09/14 a 09/15.

Como não ocorreu dependência espacial, para a estimação e construção do mapa foi utilizado o interpolador IDW. Percebe-se que os clientes com maior venda tendem a se concentrar na região inferior esquerda do mapa. Já os clientes com menor venda estão mais dispersos na região, como mostrado na Figura 6.

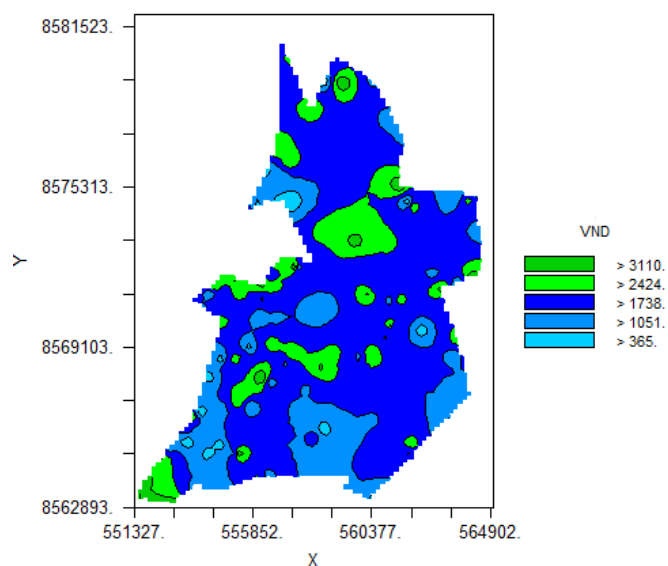


Figura 6: Mapa por interpolação IDW dos clientes varejistas de um atacadista de Uberlândia-MG, no estado da Bahia, no período de 09/14 a 09/15.

## 4.2 Dados temporais

Na análise descritiva dos dados históricos (Tabela 2), verificou-se também que existem valores semelhantes entre a média e a mediana e um coeficiente de variação moderado de 40,30%.

Tabela 2: Estatística descritiva dos dados históricos das vendas totais mensais de clientes varejistas de um atacadista de Uberlândia-MG, no estado da Bahia, no período de 01/10 a 09/15.

| Medidas       | Valores   |
|---------------|-----------|
| Mínimo        | 629.130   |
| Máximo        | 2.960.494 |
| Média         | 1.639.957 |
| Mediana       | 1.549.902 |
| Desvio Padrão | 660.880   |
| Curtose       | -1,20     |
| Assimetria    | 0,29      |
| CV            | 40,30%    |

É possível notar ao observar a Figura 7, a ausência de outliers nos dados apresentados.

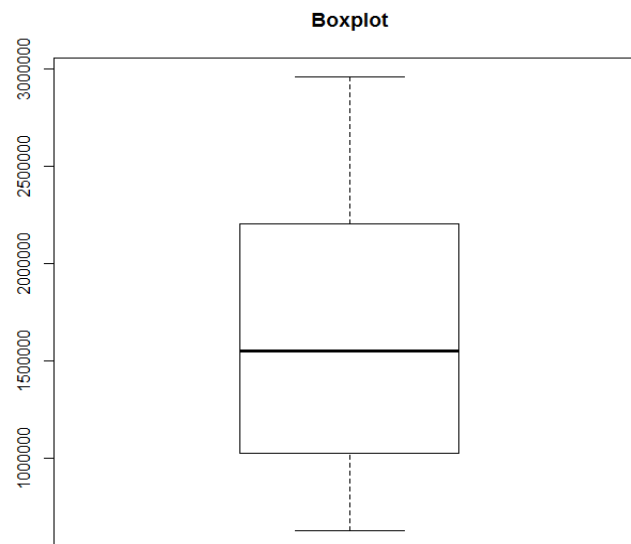


Figura 7: Boxplot das vendas totais mensais de clientes varejistas de um atacadista de Uberlândia-MG, no estado da Bahia, no período de 01/10 a 09/15.

O histograma da Figura 8 mostra uma assimetria à direita e uma curtose platicúrtica, ou seja, com caudas mais curtas e menor probabilidade que a normal de ter outliers.

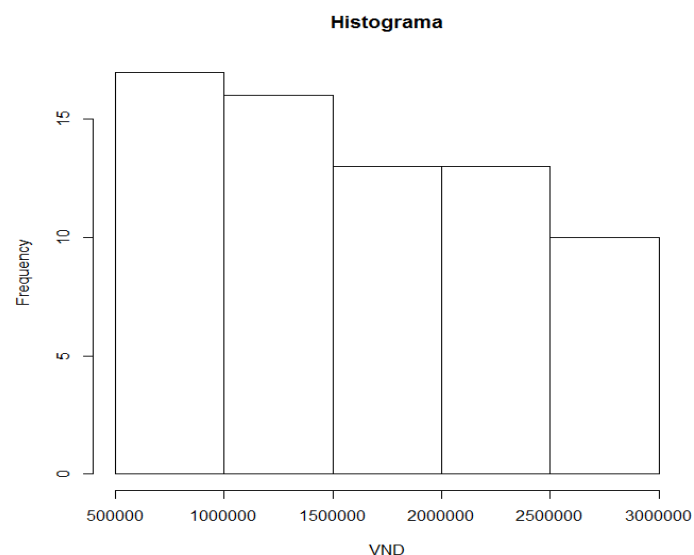


Figura 8: Histograma das vendas totais mensais de clientes varejistas de um atacadista de Uberlândia-MG, no estado da Bahia, no período de 01/10 a 09/15.

Após a análise descritiva geral, foi feito um gráfico de linhas para visualizar o comportamento da série, quanto a presença de tendência ou sazonalidade. Assim, por

meio da Figura 9 é possível notar uma tendência crescente e sazonalidade mais acentuada a partir de meados de 2013.

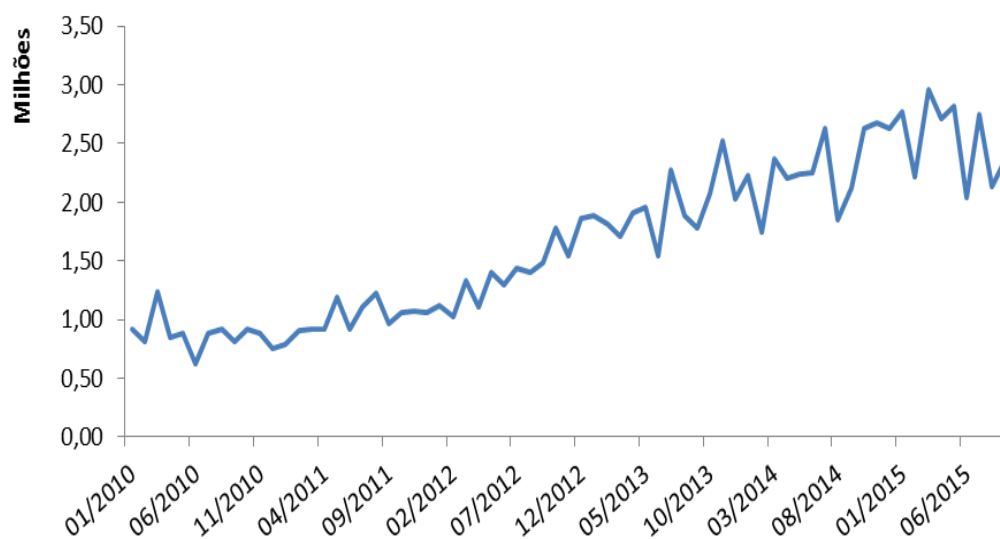


Figura 9: Comportamento da série temporal das vendas totais mensais de clientes varejistas de um atacadista de Uberlândia-MG, no estado da Bahia, no período de 01/10 a 09/15.

De acordo com a Tabela 3, concluiu-se que a série é não estacionária (p valor maior que a significância de 5%) e há a presença de tendência (p valor menor que a significância).

Tabela 3: Testes para avaliar presença de estacionariedade e tendência das vendas totais mensais de clientes varejistas de um atacadista de Uberlândia-MG, no estado da Bahia, no período de 01/10 a 09/15.

| Testes        | p-valor  |
|---------------|----------|
| Dickey-Fuller | 0,2727   |
| Cox Stuart    | < 0,0001 |

Após a análise descritiva, foram testados os ajustes de vários modelos para explicar a série. Dentre os modelos, foram utilizados os dois com melhor ajuste, Holt-Winters aditivo e SARIMA (0,1,1) (1,1,0). Os dois modelos tiveram um  $r^2$  relativamente alto, com o modelo aditivo de 0,923 e o SARIMA de 0,859, o que demonstra um bom ajuste (Tabela 4).

Para analisar se os resíduos são i.i.d. (independentes e identicamente distribuídos) foi feito o teste Ljung-Box, e nos dois modelos não temos indício para rejeitar a hipótese nula, concluindo que os resíduos são i.i.d., pois os p-valores foram maiores que 0,05. Graficamente, pode-se ver na Figura 11.

Tabela 4: Comparativo entre os modelos de séries temporais das vendas totais mensais de clientes varejistas de um atacadista de Uberlândia-MG, no estado da Bahia, no período de 01/10 a 09/15.

| Modelos              | r <sup>2</sup> | * Ljung-Box | EAM        | EQM        |
|----------------------|----------------|-------------|------------|------------|
| Holt-Winters Aditivo | 0,923          | 0,420       | 149.263,73 | 186.552,83 |
| SARIMA               | 0,859          | 0,442       | 172.734,31 | 229.309,54 |

\* p-valor

Além do critério do r<sup>2</sup>, o EAM (erro absoluto médio) e o EQM (erro quadrático médio) foram utilizados para definir qual modelo foi mais adequado para a série. Como o modelo Holt-Winters Aditivo foi o que apresentou os menores erros, ele foi eleito como sendo o modelo adequado para esta série.

Para os parâmetros de nível, tendência e sazonalidade, este modelo apresentou as estimativas 0,302; 0,004 e 0,001, respectivamente. De acordo com a Tabela 5, em outubro de 2015 o modelo selecionado apresentou erro de estimativa de 1,52%, enquanto o SARIMA errou em 4,67%.

Tabela 5: Predição dos modelos ajustados de séries temporais das vendas totais mensais de clientes varejistas de um atacadista de Uberlândia-MG, no estado da Bahia, no período de 01/10 a 09/15.

| Mês    | Observado        | Holt-Winter<br>aditivo | Sarima    |
|--------|------------------|------------------------|-----------|
| Jan/15 | 2.770.643        | 2.673.805              | 2.682.604 |
| Fev/15 | 2.215.891        | 2.502.799              | 2.414.530 |
| Mar/15 | 2.960.494        | 2.747.298              | 2.643.319 |
| Abr/15 | 2.714.882        | 2.672.763              | 2.747.385 |
| Mai/15 | 2.818.721        | 2.814.907              | 2.779.536 |
| Jun/15 | 2.040.748        | 2.511.526              | 2.608.542 |
| Jul/15 | 2.748.331        | 2.768.260              | 2.962.153 |
| Ago/15 | 2.131.414        | 2.482.630              | 2.290.790 |
| Set/15 | 2.369.871        | 2.394.371              | 2.339.394 |
| Out/15 | <b>2.632.301</b> | 2.672.405              | 2.755.266 |
| Nov/15 |                  | 2.719.447              | 2.988.108 |
| Dez/15 |                  | 2.641.259              | 2.736.993 |

O comportamento dos modelos, em relação aos valores observados, pode ser visto graficamente na Figura 10.

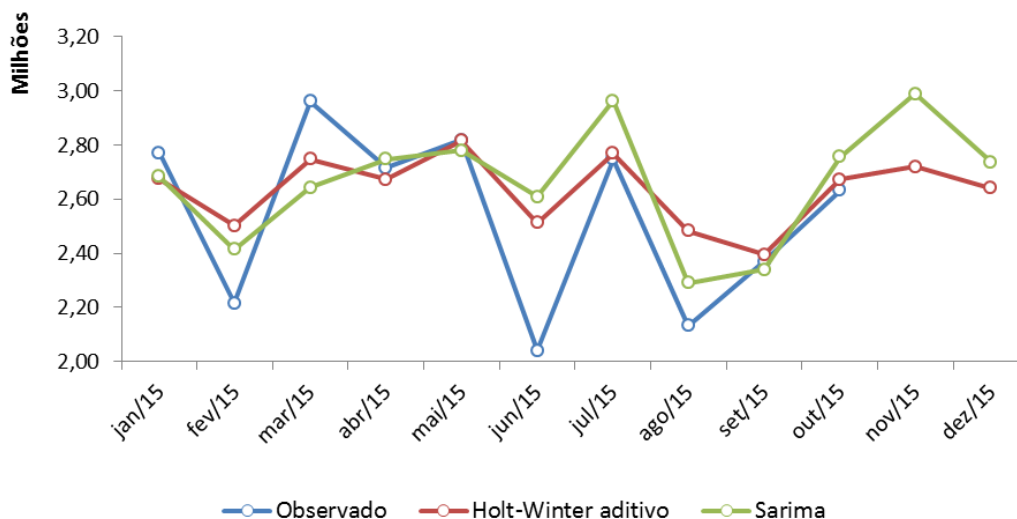


Figura 10: Comparativo dos modelos ajustados e valores observados das vendas totais mensais de clientes varejistas de um atacadista de Uberlândia-MG, no estado da Bahia, no período de 01/10 a 09/15.

Comparando os valores observados com os valores dos modelos, pela Figura 10, verifica-se que na maioria das vezes o modelo Holt-Winter Aditivo se aproxima mais dos valores observados. No mês de junho percebe-se uma alta flutuação, pois houve definições da empresa que afetaram as vendas, como por exemplo, estoque, preço, e outras.

A partir de outubro, os dois modelos projetaram valores bem diferentes, e o que se espera acontecer é que a venda se mantenha em torno de 2,6 milhões, pois a empresa enfrenta aumento de preços dos produtos da indústria, rupturas no estoque, alteração do local de armazenagem dos produtos, etc.

Completando a análise da adequacidade do modelo selecionado, na Figura 11, são apresentados os gráficos da função autocorrelação e autocorrelação parcial dos resíduos e, em conformidade, com o teste de Ljung-Box, nota-se autocorrelação e autocorrelação parcial dos resíduos não significativa.

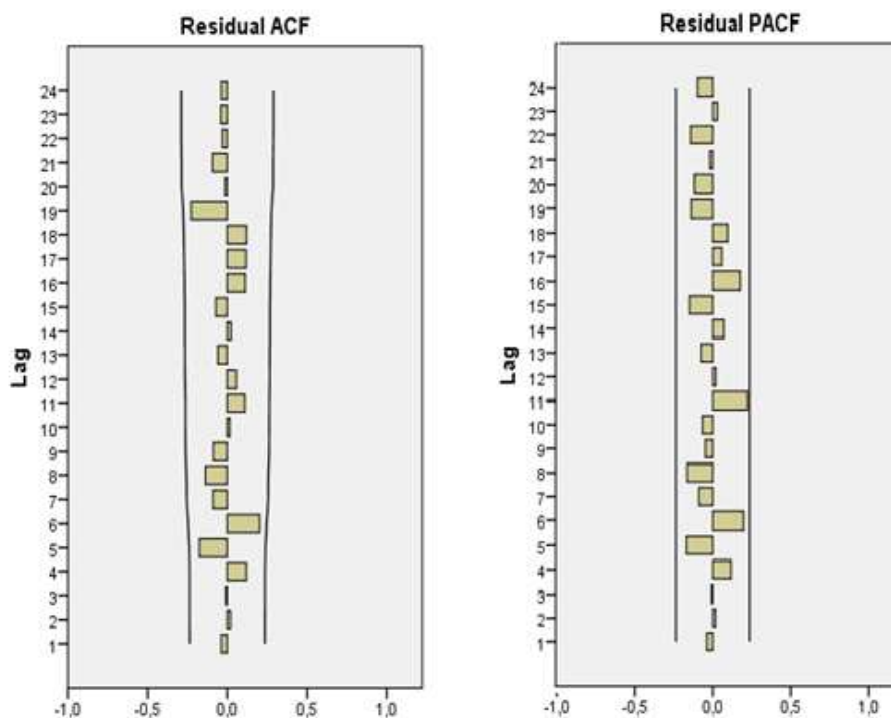


Figura 11: Autocorrelação e autocorrelação parcial dos resíduos para o modelo Holt-Winters das vendas totais mensais de clientes varejistas de um atacadista de Uberlândia-MG, no estado da Bahia, no período de 01/10 a 09/15.

Após a definição do modelo, foi gerado um *forecast* até outubro de 2016 (Figura 12). Percebe-se que o modelo gerou um intervalo de confiança pequeno, o que comprova que ele explica muito bem o comportamento dos dados.

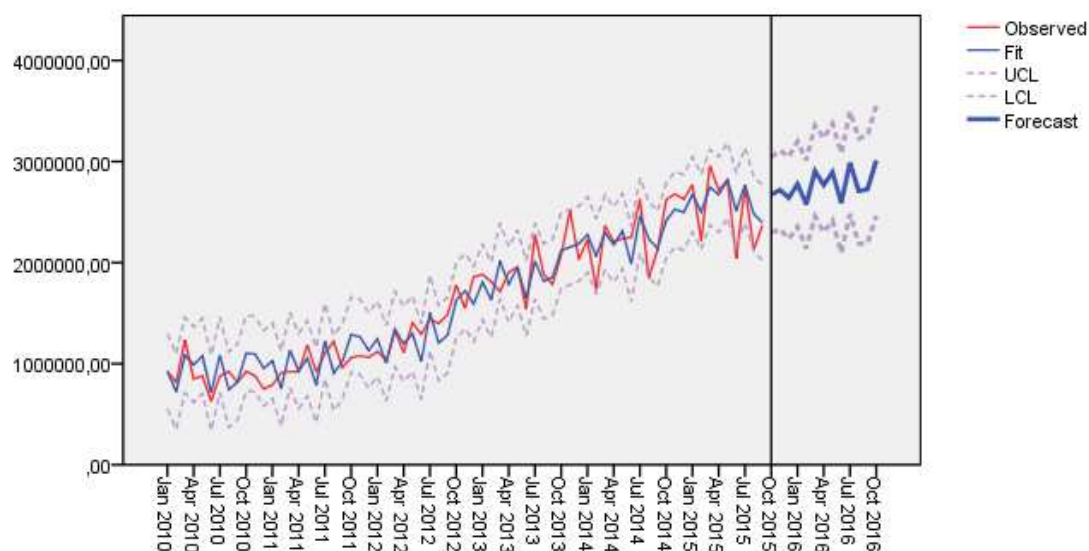


Figura 12: Série observada, modelo ajustado, *forecast* e limites do intervalo de confiança de 95%, das vendas totais mensais de clientes varejistas de um atacadista de Uberlândia-MG, no estado da Bahia, no período de 01/10 a 09/15.

Modelos de projeção possuem grande importância, pois a partir deles a empresa pode se preparar com antecedência para atingir estes resultados, planejando estoque, investimentos em sazões, diminuindo custos desnecessários, enfim, ter mais assertividade.

Em relação à análise espacial, que também contribui com o planejamento das empresas permitindo localizar e visualizar os acontecimentos e, conseqüentemente, pode ser utilizada em intervenções localizadas, recomenda-se, em trabalhos futuros, a investigação por outras técnicas, como por exemplo, o índice de Moran, buscando-se caracterizar o padrão da distribuição espacial das vendas dos varejistas nesta e em outras regiões de atuação do Atacadista.

# Capítulo 5

## Conclusão

De acordo com a análise geoestatística, para estes dados, não foi verificada dependência espacial.

O mapa gerado por IDW (interpolador não espacial) indicou que as maiores vendas se concentravam mais na parte sul da Bahia.

O comportamento temporal das vendas apresentou tendência crescente e sazonalidade que podem ser descritas pelo modelo de Holt-Winters aditivo.

## Referências

BALLOU, R. H. **Gerenciamento da cadeia de suprimentos: planejamento, organização e logística empresarial**. 4ª ed. Porto Alegre: Bookmann, 2001.

BARROS, Mônica. **Processos estocásticos**. Papel Virtual Editora, 2009.

BOX, G. E. P., JENKINS, G. M., AND REINSEL, G. C. **Time Series Analysis, Forecasting and Control**, 3rd ed. Prentice Hall, Englewood Cliffs, NJ, 2011.

BOX, G. E. P.; JENKINS, G. M. **Time series analysis: forecasting and control, revised ed**. Holden-Day, 1976.

BRAGA, L. P. V. **Geoestatística e aplicações**. In: Simpósio Brasileiro de Probabilidade e Estatística, 9., São Paulo, 1990. Resumos, São Paulo: IME-USP, 1990, p.36.

CHRISTENSEN, O. F.; RIBEIRO JR, P. J. **geoRglm-a package for generalised linear spatial models**. **R News**, v. 2, n. 2, p. 26-28, 2002.

COBRA, M. **Administração de marketing no Brasil**. Elsevier Brasil, 2009.

DAS, VARIABILIDADE ESPACIAL. **SEÇÃO IV-FERTILIDADE DO SOLO E NUTRIÇÃO DE PLANTAS**. Revista Brasileira de Ciência do Solo, v. 27, n. 6, p. 1013-1020, 2003.

GS+. **Geoestatistical for environmental sciences: version 5.0**. Michigan: Gamma Design Software, 2000.

GUIMARÃES, E.C. **Variabilidade especial de atributos de um Latossolo Vermelho-Escuro Textura Argiloso da Região do Cerrado submetido ao plantio direto e ao plantio convencional**. 2000. 85p. Tese (Doutorado em Engenharia Agrícola) – Universidade Estadual de Campinas, Campinas.

GUIMARÃES, E. C.; **Geoestatística básica e aplicada**. Uberlândia, Universidade Federal de Uberlândia, 2004.

JAKOB, A. A. E.; YOUNG, A. F. **O uso de métodos de interpolação espacial de dados nas análises sociodemográficas**. Trabalho apresentado no XV Encontro Nacional de Estudos Populacionais, ABEP, realizado em Caxambu, MG, Brasil, 2006.

KEIM, D. A. et al. (Ed.). **Mastering the information age-solving problems with visual analytics**. Florian Mansmann, 2010.

KOTLER, P. **Administração de marketing: análise, planejamento, implementação e controle**. 5.ed. São Paulo: Atlas, 1998.

MACHADO, L. O. de; LANA, A. M. Q.; QUINTÃO, R. M.; **Variabilidade espacial de atributos químicos do solo em áreas sob sistema plantio convencional**. Revista Brasileira de Ciência do Solo, v. 31, n. 3, p. 591-599, 2007.

MORETTIN, P. A.; TOLOI, C. M. C. **Modelos de função de transferência**. São Paulo: ABE, 1989.

MATTAR, F. et al. **Pesquisa de Marketing, 7ª Edição**. Elsevier Brasil, 1999.

MORETTIN, P.A. & TOLOI, C.M. **Análise de Séries Temporais**. Edgar Blucher Ltda. São Paulo, 2004.

PARENTE, J. **Varejo no Brasil, gestão e estratégia**. São Paulo: Atlas, 2000. 388p.

PELLEGRINI, F. R. & FOGLIATTO, F. S. **Estudo comparativo entre os modelos de Winters e de Box-Jenkins para previsão de demanda sazonal**. Revista Produto & Produção, Porto Alegre, v. 4, n. especial, p. 72-85, abr. 2000.

RIBEIRO, Jr, P. J.; DIGGLE, P. J. **geoR: a package for geostatistical analysis**, R News 1/2: 15–18. 2001 Available from: <http://www.rproject.org/doc/Rnews>.

SÁFADI, T.; **Uso de séries temporais na análise de vazão de água na represa de furnas.** Ciência e Agrotecnologia, v. 28, n. 1, p. 142-148, 2004.

SAMOHYL, R.W; ROCHA, R. & MATTOS, V.L.D de. **Utilização do método de Holt-Winters para previsão do leite entregue às indústrias catarinenses.** XXI Encontro Nacional de Engenharia de Produção. Salvador, 2001.

SILVA, V. R.; REICHERT, J. M.; STORCK, L. ; FEIJO, S. **Variabilidade especial das características químicas do solo e produtividade de milho em um Argissolo Vermelho-Amarelo distrófico arênico.** In: Revista Brasileira de Ciência do Solo, Viçosa, 2003.

SLACK, N. & CHAMBERS, Stuart & JOHNSTON, R.; **Administração da Produção.** 2° a ed. São Paulo: Atlas. 2002.

SOARES, A.; **Geoestatística para as ciências da terra e do ambiente.** Instituto Superior Técnico, 2000.

SOLOMON, M. R.; **Comportamento Do Consumidor, O.** Artmed, 2007.

SPSS, IBM. "IBM SPSS statistics base 20." **Chicago, IL: SPSS Inc.,** 2011.

TRANGMAR, B.B.; YOST, R.S.; WADE, M.K.; UEHARA, G. **Applications of geostatistics to spatial studies of soil properties.** Advances in Agronomy, v.38, p.45-94, 1985.

TUBINO, D. F.; **Manual de planejamento e controle da produção.** 2. ed. São Paulo: Atlas, 2000.

VIEIRA, S. R. et al. **Geoestatística em estudos de variabilidade espacial do solo.** Tópicos em ciência do solo. Viçosa: Sociedade Brasileira de Ciência do Solo, v. 1, p. 1-54, 2000.

VIEIRA, S. R.; HATFIELD, J. L.; NIELSEN, D. R.; BIGGAR, J. W. **Geostatistical theory and application to variability of some agronomical properties**. University of California, Division of Agricultural Sciences, 1983.