



# ConBRepro

XIII CONGRESSO BRASILEIRO DE ENGENHARIA DE PRODUÇÃO



IA nas Engenharias

29 nov. a 01  
de dezembro 2023

## Previsão de vendas aplicada em base de dados de uma empresa do segmento de varejo do e-commerce

**Paulo Sergio da Silva**

Programa de Pós-Graduação em Engenharia de Produção - UFPR

**Filipe Eduardo Martins Guedes**

Programa de Pós-Graduação em Engenharia de Produção - UFPR

**Mariana Kleina**

Programa de Pós-Graduação em Engenharia de Produção - UFPR

**Izabel Cristina Zattar**

Programa de Pós-Graduação em Engenharia de Produção - UFPR

**Robson Seleme**

Programa de Pós-Graduação em Engenharia de Produção - UFPR

**Resumo:** Analisar o histórico de uma empresa no tocante a suas vendas pode ser uma grande vantagem para uma organização. Entender, a partir do comportamento passado, as previsões de demanda, gera uma fonte ímpar de informações que, se estrategicamente utilizadas no tocante a compras e estoque, tornará a empresa mais forte dentro do mercado em que atua. O presente artigo busca, a partir da base estudada e dos métodos aplicados, descobrir qual o melhor método para previsão de demanda de uma empresa do segmento logístico. Para tanto o software R foi utilizado como ferramenta de trabalho, onde todos os passos para a aplicação dos métodos das médias móveis simples, suavização exponencial simples e ARIMA são demonstrados. Ao final, analisando a menor Raiz do Erro Médio Quadrático, foi possível identificar que o modelo ARIMA foi a melhor técnica de previsão de demanda, levando em consideração a base de dados escolhida e o tratamento inicial dado a ela.

**Palavras-chave:** Previsão de vendas, Médias móveis simples, Suavização exponencial simples, ARIMA, vendas agregadas.

## Sales forecasting applied to a retail e-commerce company's database

**Abstract:** Analyzing a company's sales history can be a significant advantage for an organization. Understanding demand forecasts based on past behavior provides a unique source of information that, if strategically utilized in purchasing and inventory management, can strengthen the company's position within its market. This article aims to determine, based on the studied dataset and applied methods, which is the best method for demand forecasting in a logistics company. To achieve this, the R software was employed as a working tool, showcasing the steps involved in applying the methods of simple moving averages, simple exponential smoothing, and ARIMA. In conclusion, by analyzing the lowest Root Mean Squared Error, it was possible to identify that the ARIMA model was

the most effective demand forecasting technique, considering the selected dataset and the initial data preprocessing.

**Keywords:** Sales forecasting, Simple moving averages, Simple exponential smoothing, ARIMA, Aggregate sales.

## 1. Introdução

Em um mercado cada vez mais acirrado, onde a competitividade aumenta a cada dia, aliada com a pressão pela redução de custos, faz com que as empresas busquem maneiras de se diferenciarem no mercado (FERRO *et al.*, 2019).

A Logística e o Marketing são considerados grandes armas para as organizações. Não adianta vender se a organização não tiver a capacidade de entregar e não adianta ter a capacidade de entregar sendo que não há vendas (ZHENG *et al.*, 2020; WANG, 2024).

Uma forma de equilibrar as questões de oferta e demanda vivenciadas pelas empresas é saber da forma mais assertiva a previsão da sua demanda. Como o fator tempo pode muitas vezes impactar no planejamento de previsões futuras, a tecnologia auxilia os gestores logísticos a tornarem este processo mais célere.

A utilização da previsão de demanda pode ser um diferencial competitivo para as organizações, pois, com a aplicação correta desta estratégia, a tendência é que seus estoques não fiquem sobrecarregados ou com falta de produtos, o que resultaria na falha do atendimento aos clientes e a conseqüente queda no nível de serviço. As previsões de demanda devem fazer parte do planejamento estratégico de qualquer organização que almeje um equilíbrio entre a oferta e demanda (FERRO *et al.*, 2019).

Diante do exposto, o objetivo geral deste estudo é responder a seguinte pergunta de pesquisa: Dentre os métodos estudados com base na série escolhida e com os tratamentos primários executados, qual o melhor método de previsão de demanda?

Visando atender o objetivo geral do artigo, foram aplicados três métodos de previsão de demanda por meio do *software* R, onde foi possível perceber o desempenho de cada método aplicado e sua Raiz do Erro Médio Quadrático, fator determinante para a escolha do melhor método de acordo com a base de dados escolhida e seu tratamento primário.

O artigo está dividido em 5 seções: 1) Introdução, 2) Metodologia, 3) Referencial teórico, 4) Resultados e análises e 5) Conclusão.

## 2. Metodologia

### 2.1 Base de dados

Os dados para fundamentação deste estudo foram obtidos no site <http://www.kaggle.com>, uma plataforma de competições e uma comunidade online composta por cientistas de dados e profissionais de aprendizado de máquina associados à *Google*.

O conjunto de dados selecionado, denominado “*Dataco Smart Supply Chain for Big Data Analysis*,” é uma coleção de informações de cadeias de suprimentos fornecida pela empresa *DataCo Global*. Ele abrange todas as vendas da empresa durante o período de 01/01/2015 a 31/01/2018, contendo um total de 180.519 registros.

## 2.2 Tratamentos dos dados

Dado que o propósito deste estudo é apresentar a construção de uma previsão de vendas para um período específico, foram aplicadas algumas operações de tratamento nos dados, garantindo que o conjunto de dados selecionado fornecesse os resultados desejados.

Inicialmente, houve a aplicação de um filtro no campo “*Order.Status*” para selecionar apenas as vendas com o status “*COMPLETE*” (completa), uma vez que os demais status não representavam vendas confirmadas, ou seja, vendas que foram concluídas com a entrega dos produtos aos clientes. A Tabela 1 exibe uma sumarização dos diferentes status encontrados no conjunto de dados, juntamente com suas quantidades correspondentes e o valor total de vendas associado a cada um deles.

**Tabela 1 – Sumarização do conjunto de dados – *Order.Status***

<i>Order.Status</i>	Qtde de Registros	Valor das Vendas
<i>CANCELED</i>	3.692	744.370,40
<i>CLOSED</i>	19.616	4.022.624,20
<i>COMPLETE</i>	59.491	12.095.315,00
<i>ON_HOLD</i>	9.804	1.981.542,70
<i>PAYMENT_REVIEW</i>	1.893	383.653,70
<i>PENDING</i>	20.227	4.120.532,90
<i>PENDING_PAYMENT</i>	39.832	8.106.697,60
<i>PROCESSING</i>	21.902	4.504.063,70
<i>SUSPECTED_FRAUD</i>	4.062	825.935,00

**Fonte: Os autores (2023)**

Como resultado da filtragem do campo “*Order.Status*”, obteve-se um total de 59.491 registros contendo vendas confirmadas, que serão usados nos cálculos para os métodos de previsão de vendas. Em seguida, foi necessário realizar um tratamento no campo “*order.date..DateOrders*” e convertê-lo para o formato de data “YYYY-MM-DD”.

Isso permitiu agrupar os valores das vendas por datas únicas, resumindo o campo “*Sales*” de acordo com o campo “*order.date..DateOrders*”. Isso resultou em um novo conjunto de dados contendo apenas 1.127 registros, representando o total de vendas (em valor monetário) para cada dia do período do conjunto de dados.

Para criar uma série temporal adequada para a aplicação de diversos métodos de previsão, o novo conjunto de dados precisou ser reclassificado em ordem crescente, do mais antigo para o mais recente, com base no campo “*order.date..DateOrders*”.

Após a conclusão de todos os tratamentos de dados, a seleção final de campos incluiu as variáveis “*order.date..DateOrders*” e “*Sales*”. Esse novo conjunto de dados, agora denominado “vendas”, foi construído para representar o volume diário de valor vendas da empresa. A Figura 1 ilustra essa série temporal em um gráfico.

Figura 1 – Série Temporal (vendas) – Novo conjunto de dados tratados

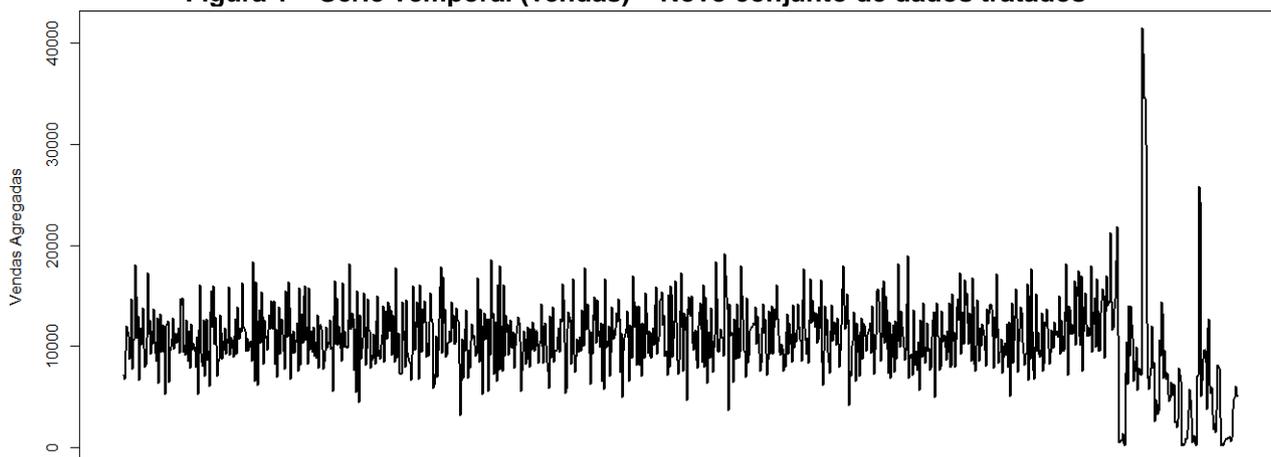


Figura 1 – Os autores (2023)

## 2.3 Métodos de Previsão e Linguagem R

Para o desenvolvimento de todos os algoritmos de tratamento de dados e criação dos métodos de previsão, utilizou-se a linguagem de programação R, que não apenas é uma linguagem de programação, mas também um ambiente rico em recursos para análise estatística e geração de gráficos.

No *software* R, para o método ARIMA, foi utilizada a biblioteca *forecast* e a função *auto.arima()*. Para o método de Suavização Exponencial Simples foi utilizada a biblioteca *forecast* e a função *ses()*. Para o método de Médias Móveis Simples não foi utilizada nenhuma função pronta do R, sendo o método implementado pelos autores.

## 3. Referencial teórico

Existem diversos métodos de se prever as demandas partindo de uma série temporal de dados. A seguir serão apresentados os métodos da Média Móvel Simples, Suavização Exponencial Simples e ARIMA.

### 3.1 Média Móvel Simples

A Média Móvel Simples (MMS) fornece uma previsão do valor médio de uma série temporal. Estas previsões podem ser obtidas por meio da média móvel simples caso a série estudada não tenha qualquer tendência perceptível ou sazonalidade. A Média Móvel Simples, apresentada na Equação 1, calcula o valor previsto por meio dos  $k$  valores recentes de uma série de dados (SWARI *et al.*, 2021).

A equação para o cálculo da Média Móvel Simples é apresentada na Equação 1.

$$MMS_k = \frac{1}{k} \sum_{i=1}^k R_i \quad (1)$$

Em que  $R_i$  é o valor real no instante de tempo passado  $i$ . E o parâmetro  $k$ , é muito importante, pois define a quantidade de dados passados a ser usado na previsão, e ele é dependente da correlação desses dados.

Prever a demanda é um método que as empresas utilizam para que possam ajustar seus estoques e garantir um suprimento eficiente, baseado em consumos passados. Quando uma organização realiza uma previsão de demanda, ela está procurando entender os comportamentos passados para poder planejar suas vendas futuras (ACKEMANN; SELLITO, 2022).

### 3.2 Suavização Exponencial Simples

O método de Suavização Exponencial Simples (SES) exige que a série de dados tenha duas particularidades: não pode apresentar tendências nem sazonalidade, ou seja, o método é indicado para séries estacionárias (HYNDMAN; ATHANASOPOULOS, 2018).

A expressão para calcular a Suavização Exponencial Simples é descrita pela Equação 2.

$$P_t = \alpha R_{t-1} + (1 - \alpha)P_{t-1} \quad (2)$$

Na Tabela 2 consta a descrição dos itens da equação do método de Suavização Exponencial Simples.

**Tabela 2 – Descrição dos termos da equação de Suavização Exponencial Simples**

Termo da equação	Significado
$R_{t-1}$	Valor real no período $t - 1$
$P_{t-1}$	Valor previsto no período $t - 1$
$\alpha$	Constante de suavização, $0 \leq \alpha \leq 1$

**Fonte: Adaptado de Hyndman e Athanasopoulos (2018)**

Quanto mais o  $\alpha$  for próximo de 1, maior será o peso dado as observações recentes, ou seja, em séries com variações rápidas, o  $\alpha$  próximo de 1 deixa o método sensível a estas mudanças (HYNDMAN; ATHANASOPOULOS, 2018; CUSTÓDIO *et al.*, 2019).

### 3.3 ARIMA

O ARIMA é o acrônimo das palavras *Auto Regressive Integrated Moving Average*. O ARIMA é um modelo de média móvel integrada auto-regressiva. Este modelo possui diversas vantagens, como, por exemplo, ser extremamente versátil em relação a outros modelos de regressão, a utilização de poucos parâmetros para ajustes dos modelos, além de permitir testar hipóteses sobre os parâmetros, tratando a autocorrelação de forma explícita, ajustando os modelos de forma particionada.

Um modelo ARIMA é composto por três partes: A primeira parte define a ordem do modelo auto-regressivo, a segunda parte define se a série deve ser diferenciada, transformando-a em estacionária e a última parte que define a ordem do modelo de médias móveis. Os modelos auto-regressivos (parte AR do acrônimo ARIMA) utilizam como base as autocorrelações presentes nos dados.

Por exemplo, a quantidade de determinado produto vendida em março de determinado ano pode estar relacionada com as vendas de fevereiro, que por sua vez pode estar relacionada com as quantidades vendidas em janeiro (HYNDMAN; ATHANASOPOULOS, 2018).

Um modelo muito aplicado na variável  $R_t$  assume que a previsão pode ser feita mediante uma combinação linear dos  $p$  dados da série temporal, conforme Equação 3.

$$R_t = c + \phi_1 R_{t-1} + \phi_2 R_{t-2} + \dots + \phi_p R_{t-p} + e_t \quad (3)$$

Em que  $\phi_i$  é o  $i$ -ésimo coeficiente da combinação linear,  $c$  é uma constante,  $R_{t-i}$  são os valores reais da série temporal no período  $t - i$ , com  $i = 1, \dots, p$  e  $e_t$  é o ruído branco (média constante e igual a zero, além da variância constante).

Os modelos de médias móveis (MA do acrônimo ARIMA: *Moving Average*), utilizam valores passados das séries, considerando os erros de previsão quando de fato há uma correlação entre os resíduos.

Um modelo de médias móveis de ordem  $q$  está representado na Equação 4.

$$R_t = \mu + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_q e_{t-q} \quad (4)$$

Em que  $\theta_j$  é o  $j$ -ésimo coeficiente da combinação linear,  $\mu$  é a média da série e  $e_{t-j}$  é o erro no período  $t - j$ , com  $j = 1, \dots, q$ .

A parte I do acrônimo ARIMA nem sempre tem necessidade de ser aplicada. Esta integração consiste na diferenciação da série (valores atuais subtraídos dos valores anteriores), para que de fato ela se torne estacionária (HYNDMAN; ATHANASOPOULOS, 2018).

### 3.4 Raiz do Erro Médio Quadrático

A Raiz do Erro Médio Quadrático (REMQ) é uma métrica estatística amplamente utilizada para avaliar a precisão de modelos preditivos, representando a média da diferença entre os valores previstos e os valores reais, expressa em uma escala semelhante aos dados originais. Quanto menor o RMSE, melhor a capacidade do modelo em fazer previsões precisas (HALLACK; FILHO, 2011). A Equação 5 apresenta o cálculo do REMQ.

$$REMQ = \sqrt{\frac{\sum_{t=1}^n e_t^2}{n}} \quad (5)$$

Onde:

$t$ : instante de tempo de uma série temporal;

$R_t$ : valor real no instante de tempo  $t$ ;

$P_t$ : valor previsto no instante de tempo  $t$ ;

$e_t = R_t - P_t$ : erro observado no instante de tempo  $t$ ;

$n$ : número de períodos de previsão.

## 4. Resultados e Análises

Nesta seção são apresentados os resultados para cada método de previsão aplicado na série temporal de dados para obter a previsão de vendas.

### 4.1 Análise de estacionariedade

Para determinar quais métodos de previsão seriam adequados para a série temporal de dados, realizou-se uma análise de estacionariedade, o que torna os resultados mais confiáveis e facilita a previsão.

Dois testes foram empregados para verificar a estacionariedade. O primeiro foi o teste ADF (*Augmented Dickey Fuller*), em que a hipótese nula ( $H_0$ ) considera que a série não é estacionária, enquanto a hipótese alternativa ( $H_1$ ) considera que a série é estacionária. O resultado deste teste revelou um p-valor de 0,01, isto é p-valor  $< \alpha$ , levando à rejeição da hipótese nula e à aceitação da hipótese alternativa com  $\alpha = 0,05$ , indicando a estacionariedade da série.

O segundo teste aplicado foi o teste KPSS (*Kwiatkowski-Phillips-Schmidt-Shin*). Neste caso, a hipótese nula ( $H_0$ ) considera que a série é estacionária, enquanto a hipótese alternativa ( $H_1$ ) considera que a série não é estacionária. O resultado deste segundo teste também apresentou um p-valor de 0,01, isto é  $p\text{-valor} < \alpha$ , levando à rejeição da hipótese nula e à validação da hipótese alternativa com  $\alpha = 0,05$ , indicando que a série não é estacionária.

No ambiente de programação R, a biblioteca *tseries* foi habilitada para a realização desses testes, utilizando as funções *adf.test()* e *kps.test()*. Os resultados dos testes são ilustrados na Figura 2.

**Figura 2 – Testes de estacionariedade realizados no software R**

```
> adf.test(vendas) #se p-valor < 0.05, então a série é estacionária

Augmented Dickey-Fuller Test

data: vendas
Dickey-Fuller = -6.6641, Lag order = 10, p-value = 0.01
alternative hypothesis: stationary

warning message:
In adf.test(vendas) : p-value smaller than printed p-value
> kps.test(vendas) #se p-valor > 0.05, então a série é estacionária

KPSS Test for Level Stationarity

data: vendas
KPSS Level = 0.90843, Truncation lag parameter = 7, p-value = 0.01

warning message:
In kps.test(vendas) : p-value smaller than printed p-value
```

**Fonte: Os autores (2023)**

Observa-se uma divergência entre os resultados dos testes. Enquanto o teste *KPSS* aponta para a não estacionariedade da série, o teste *ADF* sugere estacionariedade. Por este motivo, é indicado realizar a diferenciação da série, para confirmar sua estacionariedade.

Uma abordagem direta para converter uma série temporal originalmente não estacionária em uma série estacionária é a aplicação da técnica de diferenciação. Isso envolve subtrair os valores de uma observação pelos valores da observação imediatamente anterior.

Geralmente, uma única diferenciação é suficiente para transformar uma série não estacionária em uma série estacionária. Para realizar essa diferenciação, utilizou-se a função *diff()* no software R. Em seguida, aplicaram-se novamente os dois testes de estacionariedade, desta vez na série diferenciada, que foi denominada como “*vendas\_diff*”. A Figura 3 ilustra o processo de diferenciação e os resultados dos novos testes.

**Figura 3 – Diferenciação da série no software R**

```
> vendas_diff<-diff(vendas) #primeira diferença
> adf.test(vendas_diff)

Augmented Dickey-Fuller Test

data: vendas_diff
Dickey-Fuller = -14.927, Lag order = 10, p-value = 0.01
alternative hypothesis: stationary

warning message:
In adf.test(vendas_diff) : p-value smaller than printed p-value
> kpss.test(vendas_diff)

KPSS Test for Level Stationarity

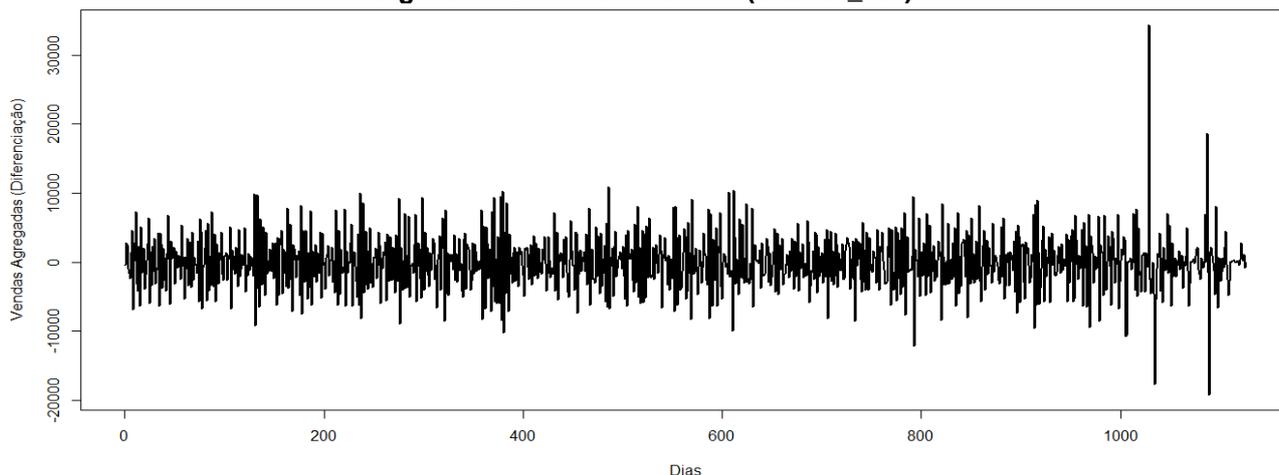
data: vendas_diff
KPSS Level = 0.0098668, Truncation lag parameter = 7, p-value = 0.1

warning message:
In kpss.test(vendas_diff) : p-value greater than printed p-value
```

Fonte: Os autores (2023)

De acordo com o resultado dos novos testes, a série diferenciada agora é estacionária, o que é mostrado na Figura 4.

**Figura 4 – Séria diferenciada (vendas\_diff)**



Fonte: Os autores (2023)

## 4.2 Análises de tendência e sazonalidade

Contudo, a fim de verificar se a série diferenciada não apresenta tendências, realizou-se uma comparação entre as séries “Vendas” e “vendas\_diff” utilizando a função *mk.test()* da biblioteca *trend*.

Para o teste de tendência, empregou-se o Teste *Mann-Kendall*, que determina a presença de tendências com base em um p-valor menor que  $\alpha = 0,05$ . Para a série “vendas”, obteve-se um p-valor de 0,02742, indicando a presença de tendências. Por outro lado, para a série “vendas\_diff”, o p-valor foi de 0,851, o que sugere a ausência de tendências. A Figura 5 apresenta uma visão comparativa dos resultados.

Figura 5 – Resultados do teste de Tendências das séries “vendas” e “vendas\_diff”

```
> mk.test(vendas)

      Mann-Kendall trend test

data: vendas
z = -4.1939, n = 1127, p-value = 2.742e-05
alternative hypothesis: true S is not equal to 0
sample estimates:
      S      varS      tau
-5.292700e+04  1.592597e+08 -8.341594e-02

> mk.test(vendas_diff)

      Mann-Kendall trend test

data: vendas_diff
z = -0.18781, n = 1126, p-value = 0.851
alternative hypothesis: true S is not equal to 0
sample estimates:
      S      varS      tau
-2.368000e+03  1.588364e+08 -3.738722e-03

> # se p-value < 0.05, então a série tem tendência
```

Fonte: Os autores (2023)

Como último passo, conduziu-se um teste de análise de sazonalidade usando o teste de Fisher, por meio da biblioteca *GeneCycle* no software R, com a função *fisher.g.test()*. Nesse teste, considerou-se que a série seria sazonal se o p-valor fosse menor que  $\alpha = 0,05$ . O resultado obtido para o p-valor foi de 0,1480991, indicando que a série não apresenta sazonalidade.

Os métodos de previsão foram aplicados à série de vendas diferenciada, garantido a estacionariedade, sem tendência e sem sazonalidade. Porém, após as previsões, os valores previstos foram revertidos para os valores originais da série, por meio da técnica oposta a da diferenciação.

### 4.3 Análises para o método de previsão Média Móvel Simples

Como visto anteriormente, a Média Móvel Simples é uma técnica de previsão que se baseia no cálculo da média dos valores anteriores em uma série temporal para estimar o valor futuro. Neste método, podem-se aplicar diferentes números de períodos passados (valores de  $k$  da Equação 1). Para esta análise, foram testados  $k = 1$ ,  $k = 2$  e  $k = 3$  períodos anteriores, resultando em  $MMS_1$ ,  $MMS_2$  e  $MMS_3$ .

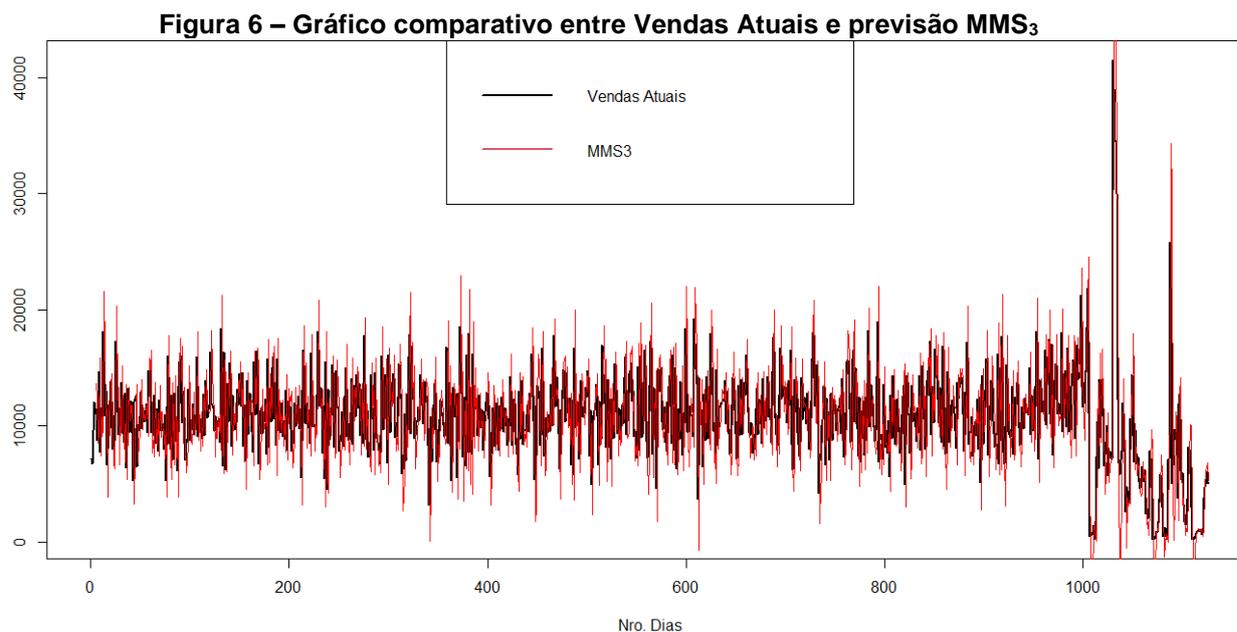
E de acordo com a Tabela 3, são apresentados os resultados do cálculo para  $MMS_1$ ,  $MMS_2$ ,  $MMS_3$ , e utilizando a função da Raiz do Erro Médio Quadrático conforme mostrado na Equação 5.

Tabela 3 – Raiz do Erro Médio Quadrático para  $MMS_1$ ,  $MMS_2$  e  $MMS_3$

Método	REMQR
$MMS_1$	6.602,897
$MMS_2$	5.211,507
$MMS_3$	4.759,966

Observa-se que a melhor opção para a previsão de vendas é obtida considerando o  $MMS_3$ , com uma Raiz do Erro Médio Quadrático de 4.759,966.

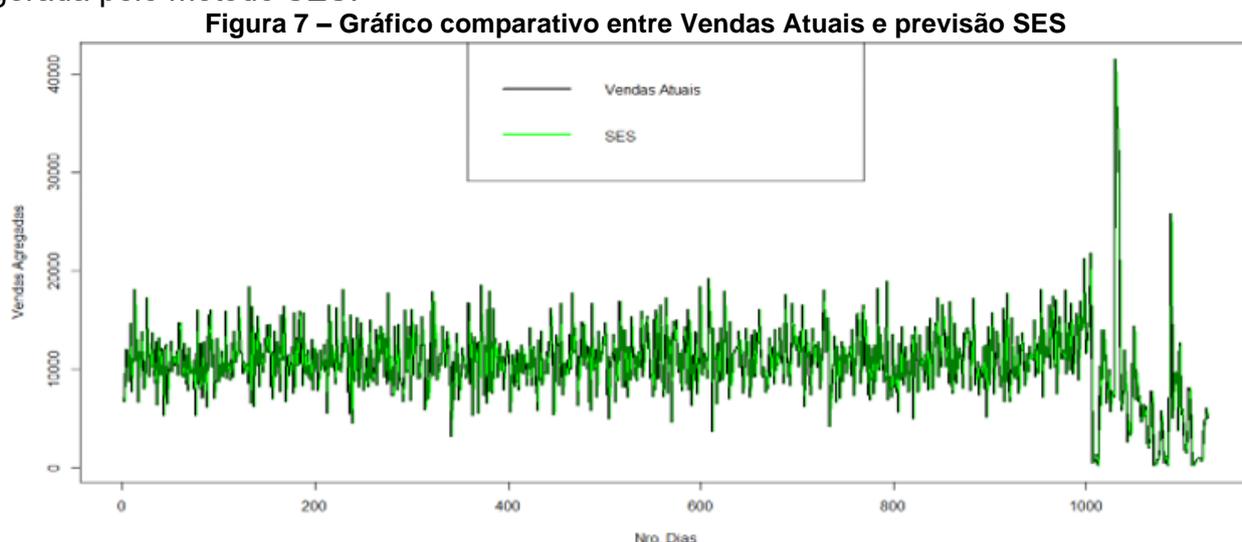
Na Figura 6 é possível observar o gráfico comparativo da série de dados original e o valor previsto pelo método  $MMS_3$ .



#### 4.4 Análises para o método de previsão Suavização Exponencial Simples

A Suavização Exponencial Simples é uma técnica de previsão que atribui maior peso aos valores mais recentes em uma série temporal. É especialmente adequada para séries estacionárias, ou seja, séries que não apresentam tendência ou sazonalidade.

Ao calcular a Raiz do Erro Médio Quadrático a esta previsão, obteve-se um valor de 3.959,192. A Figura 7 apresenta uma comparação entre as vendas reais e a previsão gerada pelo método SES.



O *software* R ajusta um valor ideal para o parâmetro  $\alpha$ , que neste caso foi 0,000100059.

## 4.5 Análises para o método de previsão ARIMA

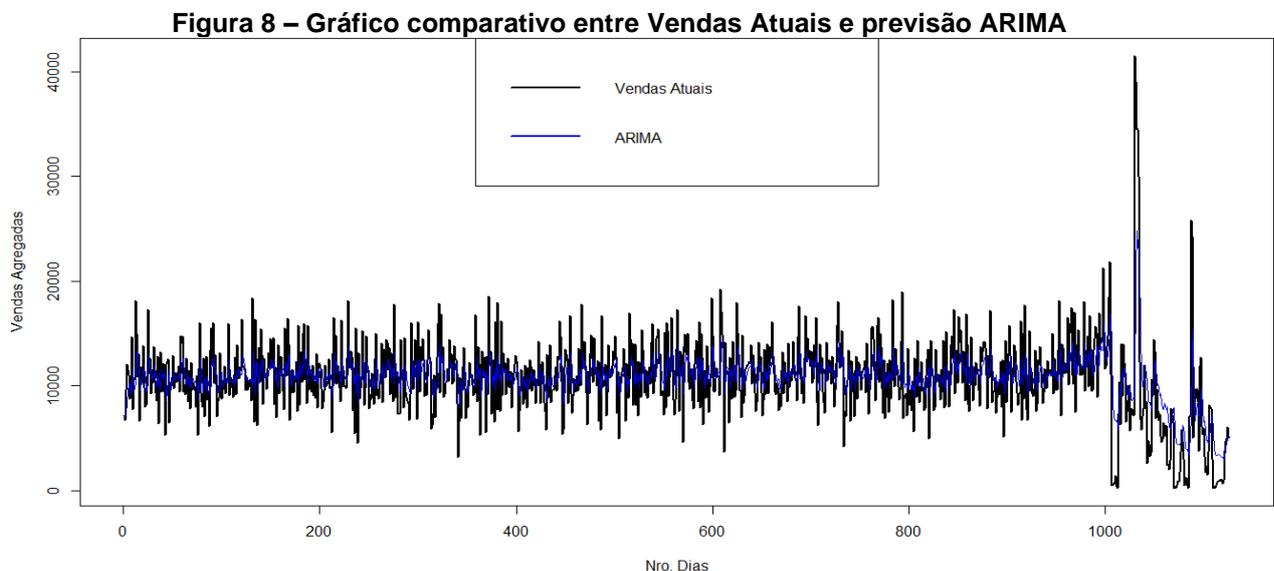
O Modelo ARIMA é um método mais sofisticado que incorpora a autocorrelação, a diferenciação e a média móvel de uma série temporal. Ele é composto por três componentes principais:

- Componente autoregressiva (AR): Refere-se à dependência da série temporal em relação a seus próprios valores passados. É representado por  $p$  no modelo ARIMA( $p, d, q$ );
- Componente de diferenciação (I): Refere-se à diferenciação da série para torná-la estacionária, se necessário. É representado por  $d$  no modelo ARIMA( $p, d, q$ );
- Componente de média móvel (MA): Refere-se à dependência dos valores passados dos erros de previsão. É representado por  $q$  no modelo ARIMA( $p, d, q$ ).

A função utilizada no *software* R realiza o ajuste automático dos parâmetros  $p, d$  e  $q$ . Para o componente  $p$ , foram selecionados 2 valores passados, para o componente  $d$ , realizou-se uma única diferenciação para tornar a série estacionária, e para o componente  $q$ , observou-se a dependência de 1 valor passado do erro de previsão, resultando em ARIMA (2, 1, 1).

A combinação linear da parte autorregressiva (AR) resultante foi  $R_t = 0,3331R_{t-1} + 0,1130R_{t-2}$  e a combinação da parte de médias móveis (MA) foi  $R_t = -0,9693e_t$ .

Ao calcular a Raiz do Erro Médio Quadrático a esta previsão, obteve-se o valor de 3.310,211. A Figura 8 ilustra a comparação entre as vendas reais e a previsão gerada pelo método ARIMA.



## 5. Conclusões

Para avaliar o desempenho dos métodos de previsão, utilizaram-se dados históricos de vendas que foram agregados e organizados sequencialmente por data, transformando-os em séries temporais. Empregaram-se três métodos de previsão adequados para séries estacionárias, sem tendência e sem sazonalidade. E para cada método, calculou-se seu desempenho por meio da Raiz do Erro Médio Quadrático para cada resultado obtido.

Os resultados das previsões foram visualizados em gráficos de séries temporais, permitindo uma comparação entre as previsões geradas pelos métodos e os valores reais de vendas.

Além disso, reportou-se o valor do erro para cada método, proporcionando uma análise quantitativa de sua eficácia na previsão da demanda de vendas.

Com base nas análises realizadas, conclui-se que o método ARIMA demonstrou ser o mais preciso para prever este conjunto de dados em formato de série temporal, resultando em uma Raiz do Erro Médio Quadrático de 3.310,211, enquanto que o método de Suavização Exponencial Simples resultou em 3.959,192, e por último o método de Média Móvel Simples com uma Raiz do Erro Médio Quadrático de 4.759,966.

De acordo com a base de dados utilizada, bem como o tratamento feito antes da aplicação dos métodos, o modelo ARIMA, por ter uma Raiz do Erro Médio Quadrático menor, se mostrou o método mais eficiente. Sendo assim, a margem de erro é menor, fazendo com que as previsões sejam as mais assertivas possíveis, impactando positivamente no nível de serviço prestado pela empresa a seus clientes.

A pesquisa limitou-se aos métodos de Média Móvel Simples, Suavização Exponencial Simples e o método ARIMA, a melhor opção versou entre os métodos estudados e aplicados por meio do *software* R.

Recomenda-se que, para trabalhos futuros, métodos de aprendizado de máquina sejam aplicados no conjunto de dados, buscando alternativas para previsão de demanda de vendas das organizações, fator extremamente importante para que uma empresa se mantenha sustentável ao longo do tempo.

## Referências

ACKERMANN, A. E. F., SELKITTO, M. A. **Métodos de previsão de demanda: uma revisão da literatura**. *Innovar*, v. 32 nº 85, p. 83-99, 2022.

FERRO, W. A. LIMA, J. D. TRENTIN, M. G. **Combinações de Métodos quantitativos na previsão de demanda de vendas de eletrodomésticos**, *Revista GEPROS, Gestão da Produção, Operações e Sistemas*, v.14, nº 5, p. 67-68, 2019.

HALLACK, R. FILHO, A.J.P., **Metodologia para análise de desempenho de simulações de sistemas convectivos na região metropolitana de São Paulo com o modelo ARPS: sensibilidade a variações com os esquemas de advecção e assimilação de dados**. *Revista Brasileira de Meteorologia*. v. 26 nº4, 2011

HYNDMANN, R.J, ATHANASOPOULOS, G. **Forecasting: principles and practice**, 2nd edition, OTexts: Melbourne, Australia, 2018.

SWARI, M.H.P., QUSYAIRI, M., MANDYARTHA, E.P., WAHANANI, H.E., **Business Intelligence System using Simple Moving Average Method (Case Study : Sales Medical Equipment at PT. Semangat Sejahtera Bersama)**. *Journal of Physics: Conference Series*. [doi:10.1088/1742-6596/1899/1/012121](https://doi.org/10.1088/1742-6596/1899/1/012121), 2021

WANG. Shubing. **Research on the Digital Marketing Strategies in the E-commerce Logistics Service Mode under the Influence of Big Data**. *Computer-Aided Design & Applications*, 21(S3), p. 39-55, 2024

ZHENG, K. ZHANG, Z. SONG, B. **E-commerce logistics distribution mode in big-data contexto: A case analysis of JD.COM**. *Industrial Marketing Management*, v. 86, p. 154-162. <https://doi.org/10.1016/j.indmarman.2019.10.009>