

OTÁVIO ARAÚJO DE ABREU

COMPARATIVO DE MODELOS DE PREVISÃO DE DEMANDA PARA CHEGADAS
DE PACIENTES NO INSTITUTO DO CÂNCER DO ESTADO DE SÃO PAULO

SÃO PAULO

2023

OTÁVIO ARAÚJO DE ABREU

COMPARATIVO DE MODELOS DE PREVISÃO DE DEMANDA PARA CHEGADAS
DE PACIENTES NO INSTITUTO DO CÂNCER DO ESTADO DE SÃO PAULO

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

SÃO PAULO

2023

OTÁVIO ARAÚJO DE ABREU

COMPARATIVO DE MODELOS DE PREVISÃO DE DEMANDA PARA CHEGADAS
DE PACIENTES NO INSTITUTO DO CÂNCER DO ESTADO DE SÃO PAULO

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

Orientador: Professor Daniel de Oliveira Mota

SÃO PAULO

2023

Dedico esse trabalho à minha mãe Andréa,
que é meu maior exemplo, sempre me
ensinando a ser uma pessoa melhor.

AGRADECIMENTOS

À minha mãe, por tudo que já fez por mim e por todos os seus ensinamentos. Obrigado por todo o esforço, tempo e dedicação para que eu e minha irmã pudéssemos evoluir pessoalmente e profissionalmente.

Ao meu pai, também engenheiro, que me transmitiu o amor pelas ciências exatas e pelo pensamento analítico. Obrigado pelo apoio e suporte fornecido durante a minha trajetória.

À minha irmã, pela amizade e companheirismo. Espero poder te transmitir o que aprendi e compartilhar minhas experiências nesse período para que sua jornada acadêmica seja a melhor possível.

A todos os meus familiares, por todos os momentos que vivemos juntos e pelas conversas, conselhos, viagens e experiências marcantes. Obrigado, em especial, aos meus familiares que já não estão mais aqui – meu vô Benedito, minha tia Luzia, minha vó Regina e minha vó Maria Helena. Espero que, onde quer que estejam, sintam-se orgulhosos de mim.

Aos meus amigos da Poli, obrigado pelo companheirismo, pelos momentos engraçados e por tudo que vivemos durante esse período de nossas vidas. Aprendi com cada um de vocês, que tornaram essa trajetória mais leve.

Aos meus amigos da SAP, Raízen e Genoa, por todos os conselhos e ensinamentos, que foram, e ainda são, fundamentais para a minha evolução profissional e pessoal.

Aos meus amigos da ETEC Presidente Vargas e do Colégio Santa Mônica, que até hoje levo comigo. Obrigado por tudo que já vivemos juntos, pelas experiências e memórias que criamos.

Ao meu professor orientador Daniel Mota, que é uma referência para mim desde que começamos a trabalhar juntos em 2019. Obrigado por todas as oportunidades, ensinamentos, conselhos e broncas que me deu durante minha trajetória na universidade. A Iniciação Científica e os projetos do MIT me abriram muitas portas, além de auxiliarem na minha evolução pessoal e profissional.

RESUMO

O Instituto do Câncer do Estado de São Paulo (ICESP) é um centro de referência no tratamento do câncer e enfrenta diariamente um alto fluxo de pacientes em busca de atendimento especializado e de qualidade. No entanto, a variação na demanda ao longo do tempo pode sobrecarregar sua capacidade operacional, resultando em elevados tempos de espera, formação de grandes filas e sobrecarga dos recursos disponíveis.

Diante desse cenário, o objetivo deste estudo é realizar um comparativo entre modelos de previsão de demanda para as chegadas de pacientes no ICESP, com o propósito de identificar o modelo mais adequado para estimar de forma precisa as taxas de chegadas de pacientes, viabilizando assim um dimensionamento eficiente dos recursos hospitalares.

Para o desenvolvimento do estudo, foi conduzida uma revisão bibliográfica que abordou os principais métodos de previsão de demanda, como suavização exponencial, modelos ARIMA e SARIMA, além dos conceitos de estatística que são utilizados nos cálculos associados. Na sequência, foi realizado o mapeamento do processo de chegadas no hospital, por meio de visitas de campo, além da análise dos dados fornecidos pelo instituto. Posteriormente, os métodos de previsão foram aplicados em relação aos dados reais e tiveram sua precisão avaliada por métricas de erro, a fim de obter o método que mais se adequa ao processo.

Dessa forma, obtemos maior assertividade sobre a demanda futura, promovendo dimensionamento mais eficiente dos recursos disponíveis. Com isso, será possível reduzir o tempo de espera dos pacientes, proporcionando uma melhoria significativa na qualidade dos serviços prestados e, conseqüentemente, aumentando a satisfação geral durante todo o processo de atendimento no ICESP.

Palavras-chave: Previsão de Demanda, Saúde, Recursos Hospitalares

ABSTRACT

The Cancer Institute of the State of São Paulo (ICESP) is a leading center for cancer treatment that faces a daily high influx of patients seeking specialized and quality care. However, the fluctuation in demand over time can strain its operational capacity, resulting in prolonged waiting times, extensive queues, and resource overload.

Given this scenario, the objective of this study is to conduct a comparative analysis of demand forecasting models for patient arrivals at ICESP. The aim is to identify the most suitable model for accurately estimating patient arrival rates, thereby enabling efficient resource allocation within the hospital.

For the study's development, a literature review was conducted covering key demand forecasting methods such as exponential smoothing, ARIMA, and SARIMA models, along with the statistical concepts used in the associated calculations. Subsequently, the arrival process at the hospital was mapped through on-site visits, complemented by an analysis of the data provided by the institute. The forecasting methods were then applied to real data, and their accuracy was measured using error metrics to determine the most fitting approach.

Thus, it is possible to achieve greater accuracy in predicting future demand, facilitating a more efficient allocation of available resources. Consequently, this approach aims to reduce patient waiting times, leading to a significant improvement in the overall quality of services and an increase in overall satisfaction throughout the entire patient care process at ICESP.

Keywords: Demand Forecasting, Health, Hospital Resources

LISTA DE FIGURAS

Figura 1 - Prédio do ICESP	23
Figura 2 - Tendência e sazonalidade para séries temporais.....	36
Figura 3 - Fluxograma dos atendimentos no térreo	58
Figura 4 - Representação 3D do fluxo de atendimento.....	59
Figura 5 - Guichês de atendimento na torre laranja	59

LISTA DE GRÁFICOS

Gráfico 1 - Série temporal (vendas de uma empresa farmacêutica)	30
Gráfico 2 - Série temporal (exemplo)	32
Gráfico 3 - Tendência (exemplo).....	32
Gráfico 4 - Sazonalidade (exemplo).....	33
Gráfico 5 - Ruído (exemplo)	33
Gráfico 6 - Autocorrelação para diferentes lags	38
Gráfico 7 - PACF	39
Gráfico 8 - Média móvel de 5 períodos	41
Gráfico 9 - Média móvel de 50 períodos	42
Gráfico 10 - Exemplo de regressão linear	43
Gráfico 11 - Suavização exponencial simples.....	45
Gráfico 12 - Suavização exponencial dupla	46
Gráfico 13 - ACF para um modelo AR(2)	50
Gráfico 14 - PACF para um modelo AR(2).....	50
Gráfico 15 - Chegadas por dia (Recepção Torre Laranja)	60
Gráfico 16 - Média de chegadas por dia da semana e por hora	61
Gráfico 17 - Média de chegadas por mês e por hora	62
Gráfico 18 - Série temporal original.....	63
Gráfico 19 - Tendência.....	63
Gráfico 20 - Sazonalidade.....	64
Gráfico 21 - Resíduos	64
Gráfico 22 - Média Móvel Simples (n = 2)	66
Gráfico 23 - Erro - Média Móvel Simples (n = 2)	66
Gráfico 24 - Histograma de erros absolutos - Média Móvel Simples (n = 2)	66
Gráfico 25 - Média Móvel Simples (n = 5)	67
Gráfico 26 - Erro - Média Móvel Simples (n = 5)	68
Gráfico 27 - Histograma de erros absolutos - Média Móvel Simples (n = 5)	68
Gráfico 28 - Regressão Linear	69
Gráfico 29 - Erro - Regressão Linear	70
Gráfico 30 - Histograma de erros absolutos - Regressão Linear	70
Gráfico 31 - Suavização Exponencial Tripla.....	73
Gráfico 32 - Erro - Suavização Exponencial Tripla.....	73

Gráfico 33 - Histograma de erros absolutos - Suavização Exponencial Tripla.....	73
Gráfico 34 - ACF	76
Gráfico 35 - PACF	76
Gráfico 36 - Aplicação do modelo ARIMA (5,0,0).....	77
Gráfico 37 - Erro - Aplicação do modelo ARIMA (5,0,0).....	77
Gráfico 38 - Histograma de erros absolutos - Aplicação do modelo ARIMA(5,0,0)..	78
Gráfico 39 - Aplicação do modelo SARIMA (5,0,0)(0,2,1,5)	80
Gráfico 40 - Erro - Aplicação do modelo SARIMA (5,0,0)(0,2,1,5)	81
Gráfico 41 - Histograma de erros absolutos - SARIMA(5,0,0)(0,2,1,5)	81

LISTA DE QUADROS

Quadro 1 - Comportamento teórico do ACF e PACF para séries estacionárias	49
Quadro 2 - Distribuição de destinos	60
Quadro 3 - Média de chegadas por dia da semana e por hora	61
Quadro 4 - Média de chegadas por mês e por hora.....	62
Quadro 5 - Erros na previsão por Média Móvel ($n = 2$)	67
Quadro 6 - Erros na previsão por Média Móvel ($n = 5$)	68
Quadro 7 - Erros na previsão por Regressão Linear.....	70
Quadro 8 - Erros na previsão por Suavização Exponencial Tripla	74
Quadro 9 - Erros na previsão pelo Modelo ARIMA	78
Quadro 10 - Erros na previsão pelo Modelo SARIMA.....	81
Quadro 11 - Erros para cada modelo de previsão utilizado	82

SUMÁRIO

1. INTRODUÇÃO.....	22
1.1. Histórico do ICESP	22
1.2. Engenharia de Produção e Medicina	23
1.3. Desafios na Gestão de Demandas no Contexto Hospitalar	24
1.4. Objetivo do Trabalho e Justificativa	25
1.5. Estrutura do Trabalho	25
2. FUNDAMENTAÇÃO TEÓRICA.....	26
2.1. Forecasting	26
2.2. Modelos de Previsão de Demanda	27
2.2.1. Modelos Explicativos.....	27
2.2.2. Modelos de Séries Temporais.....	28
2.2.3. Modelos Mistos	28
2.3. Séries Temporais.....	29
2.3.1. Panorama e aplicações.....	29
2.3.2. Componentes das séries temporais.....	30
2.3.3. Estacionariedade de séries temporais	33
2.3.4. Decomposição de séries temporais	34
2.4. Autocovariância e autocorrelação.....	36
2.4.1. Autocovariância.....	36
2.4.2. Autocorrelação	37
2.4.3. Função de Autocorrelação Parcial	38
2.5. Modelos de Previsão de Séries Temporais	40
2.5.1. Média Móvel Simples	40
2.5.2. Regressão Linear.....	42
2.5.3. Suavização Exponencial Simples	44

2.5.4. Suavização Exponencial Dupla (Holt)	45
2.5.5. Suavização Exponencial Tripla (Holt-Winters)	46
2.5.6. Modelo ARIMA	48
2.5.7. Modelo SARIMA.....	51
2.6. Erros nos Métodos de Previsão	54
2.6.1. Erro Médio (ME).....	54
2.6.2. Erro Absoluto Médio (MAE).....	55
2.6.3. Erro Percentual Absoluto Médio (MAPE)	55
2.6.4. Raiz do Erro Quadrático Médio (RMSE):	56
3. ANÁLISE DA CHEGADA DOS PACIENTES.....	58
3.1. Descrição do Processo	58
3.2. Análise dos Dados	59
4. APLICAÇÃO DOS MODELOS DE PREVISÃO	65
4.1. Média Móvel Simples.....	65
4.1.1. Período = 2	65
4.1.2. Período = 5	67
4.2. Regressão Linear.....	69
4.3. Suavização Exponencial Tripla (Holt-Winters).....	71
4.4. Modelo ARIMA.....	74
4.5. Modelo SARIMA	78
4.6. Escolha do Melhor Modelo	82
5. CONCLUSÃO.....	83
6. PRÓXIMOS PASSOS.....	84
REFERÊNCIAS BIBLIOGRÁFICAS.....	86

1. INTRODUÇÃO

1.1. Histórico do ICESP

O Instituto do Câncer do Estado de São Paulo (ICESP) é uma instituição médica de renome localizada na cidade de São Paulo, Brasil. Fundado em 2008, o ICESP é um centro de excelência dedicado à pesquisa, prevenção, diagnóstico e tratamento do câncer.

O ICESP foi concebido como uma resposta às crescentes demandas por serviços de saúde de qualidade no combate ao câncer. Sua criação foi fruto de uma colaboração entre o Governo do Estado de São Paulo, a Faculdade de Medicina da Universidade de São Paulo (FMUSP) e o Hospital das Clínicas da FMUSP. Desde sua inauguração, o ICESP tem se destacado como um dos principais centros de referência no tratamento do câncer não apenas no Brasil, mas também na América Latina.

Além disso, opera como um hospital de alta complexidade, oferecendo uma gama completa de serviços para pacientes com câncer. Sua abordagem multidisciplinar envolve equipes médicas especializadas, incluindo oncologistas, cirurgiões, radioterapeutas, e outros profissionais de saúde, que trabalham em conjunto para fornecer o melhor atendimento possível aos pacientes. Atualmente, possui 500 leitos instalados, 5 mil colaboradores e prestadores de serviços, além de 36 mil pacientes ativos (icesp.gov.br).

O instituto também desempenha um papel fundamental na pesquisa e desenvolvimento de novas terapias e abordagens para o tratamento do câncer. Seus pesquisadores conduzem estudos clínicos e experimentais em colaboração com instituições nacionais e internacionais, contribuindo significativamente para o avanço do conhecimento na área oncológica. Além disso, o ICESP está fortemente envolvido em programas de conscientização e prevenção do câncer, promovendo a importância da detecção precoce e de hábitos de vida saudáveis.

Figura 1 - Prédio do ICESP



Fonte: Google

1.2. Engenharia de Produção e Medicina

A relação entre a Engenharia de Produção e a Medicina tem se tornado cada vez mais evidente e sinérgica, impulsionada pelo avanço da análise de dados e da inteligência artificial (IA). Essa convergência está revolucionando a forma como a medicina é praticada e como os serviços de saúde são gerenciados.

A coleta e análise de dados desempenham um papel fundamental na medicina moderna. Desde registros médicos eletrônicos até imagens médicas, uma enorme quantidade de informações é gerada diariamente em hospitais e clínicas. Engenheiros de Produção, por serem especializados em otimização de processos e gerenciamento de dados, desempenham um papel crucial na estruturação, organização e interpretação desses dados. Isso permite a identificação de tendências, o monitoramento de pacientes e o desenvolvimento de protocolos de tratamento mais eficazes.

Já a inteligência artificial tem o potencial de revolucionar a medicina. Algoritmos de IA podem analisar dados médicos complexos, como imagens de

ressonância magnética, para auxiliar no diagnóstico precoce de doenças. Além disso, sistemas de IA podem prever a progressão da doença, identificar tratamentos personalizados com base no perfil genético do paciente e até mesmo auxiliar cirurgias em procedimentos complexos.

A sinergia entre a Engenharia de Produção e a Medicina está impulsionando avanços significativos na eficiência do atendimento médico, diagnóstico mais preciso e tratamentos personalizados. Essa colaboração demonstra como a aplicação da engenharia e da tecnologia pode ter um impacto positivo na qualidade de vida das pessoas.

1.3. Desafios na Gestão de Demandas no Contexto Hospitalar

A administração eficaz dos serviços de saúde em um hospital de alta complexidade, como o ICESP, enfrenta desafios significativos relacionados à gestão da demanda. A dinâmica complexa das chegadas de pacientes, combinada com a necessidade de recursos especializados, cria um ambiente desafiador para otimizar a eficiência operacional.

O ICESP, por se tratar de uma instituição de referência no tratamento contra o câncer, recebe um elevado número de pessoas diariamente, as quais vão ao hospital para realizar consultas médicas, receber tratamento intensivo, coletar sangue, ou até mesmo para internação em um dos leitos disponíveis.

Neste contexto, a oscilação nas demandas diárias, muitas vezes imprevisíveis, representa um obstáculo considerável. A variabilidade nas chegadas de pacientes pode resultar em sobrecarga dos recursos em determinados momentos e subutilização em outros, impactando diretamente na qualidade do atendimento e na experiência do paciente.

A previsão de demandas emerge como uma ferramenta crucial para enfrentar os desafios na gestão hospitalar. A capacidade de antecipar variações na demanda por serviços de saúde permite uma alocação mais eficiente de recursos, redução de tempos de espera e uma experiência aprimorada para os pacientes.

Além disso, a implementação de métodos de previsão proporciona uma base estratégica para a tomada de decisões. A antecipação das demandas permite não apenas a otimização dos recursos existentes, mas também a projeção de

necessidades futuras, contribuindo para o planejamento de investimentos e melhorias na infraestrutura.

1.4. Objetivo do Trabalho e Justificativa

Este trabalho visa explorar e comparar métodos de previsão de demanda aplicados às chegadas de pacientes no ICESP. A escolha do método mais adequado proporcionará uma base sólida para dimensionar eficientemente os recursos hospitalares, contribuindo para a redução dos tempos de espera e a melhoria geral dos serviços prestados.

A compreensão aprofundada dos padrões de demanda possibilitará ao hospital atender às necessidades dos pacientes de maneira mais eficaz. Ao mesmo tempo, a pesquisa visa contribuir para a literatura existente sobre gestão de demanda em contextos hospitalares, destacando a aplicação prática dos métodos de previsão.

1.5. Estrutura do Trabalho

O trabalho está estruturado da seguinte forma: na Seção 2, é realizada uma revisão bibliográfica abrangente dos métodos de previsão de demanda, incluindo suavização exponencial, modelos ARIMA e SARIMA, destacando sua relevância e abordando conceitos estatísticos associados a eles. Na Seção 3, descrevemos a metodologia aplicada, que inclui o mapeamento do processo de chegadas no hospital e a análise dos dados fornecidos pelo instituto. A Seção 4 apresenta a aplicação dos métodos de previsão, bem como os resultados obtidos e a análise comparativa entre os modelos. Por fim, a última seção conclui o trabalho, destacando as contribuições, limitações e possíveis direções para pesquisas futuras.

2. FUNDAMENTAÇÃO TEÓRICA

2.1. Forecasting

Forecasting, ou previsão em português, refere-se a um processo de estimativa ou projeção de eventos futuros com base em dados históricos ou informações disponíveis no presente. É uma técnica amplamente utilizada em áreas como economia, finanças, marketing, produção, gestão de estoques, entre outras, com o objetivo de tomar decisões mais assertivas e planejar ações futuras. Uma das definições clássicas de forecasting é dada por Makridakis (1998):

"Forecasting é o processo de fazer declarações sobre eventos cujos resultados reais (tipicamente) ainda não foram observados. A previsão geralmente envolve a utilização de dados históricos (como vendas passadas), projetando-os para o futuro por meio de um modelo matemático."

O forecasting é valioso porque permite às organizações prepararem-se para o futuro, evitando surpresas e tomando medidas proativas para enfrentar situações esperadas. No entanto, é importante reconhecer que nenhuma previsão é perfeita, e há sempre um grau de incerteza associado a eventos futuros. Assim, a previsão é geralmente acompanhada por medidas de intervalos de confiança ou margens de erro para expressar a incerteza nas estimativas.

Para realizar um processo de previsão, normalmente são seguidos os seguintes passos (Hyndman e Athanasopoulos, 2021):

1. Definição do Problema:

O primeiro passo envolve a definição clara do problema de previsão. Isso requer compreensão de como as previsões serão utilizadas, quem requer as previsões e como elas se encaixam na organização. Além disso, é essencial comunicar-se com todas as partes envolvidas na coleta e no uso dos dados para obter uma compreensão completa do contexto do problema de previsão.

2. Coleta de Informações:

São necessários dois tipos de informações - dados estatísticos e conhecimento especializado das pessoas envolvidas na coleta e no uso dos dados. Obter dados

históricos suficientes para ajustar um bom modelo estatístico pode ser desafiador, mas é crucial para uma previsão precisa.

3. Análise Exploratória dos Dados:

A análise exploratória dos dados é fundamental para entender os padrões e relações presentes nos dados. Isso envolve a criação de gráficos para identificar tendências, sazonalidades, ciclos de negócios e possíveis outliers. Essa análise preliminar ajudará a escolher o modelo de previsão adequado.

4. Escolha e Ajuste de Modelos:

Com base na análise exploratória e nas informações coletadas, o próximo passo é escolher o modelo de previsão mais adequado. Isso pode incluir a comparação de diferentes modelos potenciais. Cada modelo é construído com base em suposições e envolve parâmetros que devem ser estimados usando os dados históricos disponíveis.

5. Utilização e Avaliação do Modelo de Previsão:

Uma vez selecionado um modelo e estimados seus parâmetros, o modelo é usado para fazer previsões futuras. Sua avaliação de desempenho só pode ser feita após o período de previsão ter ocorrido, comparando as previsões com os valores reais.

Esses passos fornecem uma estrutura geral para a realização de previsões, mas a complexidade do processo pode variar dependendo do contexto específico e da natureza dos dados envolvidos.

2.2. Modelos de Previsão de Demanda

De acordo com o livro "Forecasting: Principles and Practice" de Rob J. Hyndman e George Athanasopoulos, existem três principais tipos de modelos de previsão de demanda:

2.2.1. Modelos Explicativos

Os modelos explicativos, também conhecidos como modelos causais, buscam explicar a relação entre a variável dependente (demanda) e as variáveis independentes (fatores explicativos). Esses modelos utilizam técnicas estatísticas, como regressão linear, para identificar e quantificar a influência das variáveis preditoras na variável de interesse. Os modelos explicativos são adequados quando há um entendimento claro das relações causais entre as variáveis e quando se deseja compreender os impactos individuais de cada variável preditora.

$$DE = f(\textit{temperatura}, \textit{população}, \textit{hora do dia}, \textit{erro})$$

Fonte: Adaptado de Hyndman, R. J., & Athanasopoulos, G. (2021)

2.2.2. Modelos de Séries Temporais

Os modelos de séries temporais são usados quando se trabalha com dados que são observados sequencialmente ao longo do tempo, ou seja, há uma dependência temporal nas observações. Esses modelos são projetados especificamente para lidar com padrões temporais, como tendências, sazonalidades e ciclos. Alguns dos modelos de séries temporais comuns incluem a média móvel, suavização exponencial, modelos ARIMA (AutoRegressive Integrated Moving Average) e modelos sazonais como SARIMA (Seasonal AutoRegressive Integrated Moving Average). Esses modelos são úteis quando não há informações ou variáveis explicativas adicionais disponíveis, ou quando a natureza do comportamento da série é mais relevante para a previsão.

$$DE_{t+1} = f(DE_t, DE_{t-1}, \dots, \textit{erro})$$

Fonte: Hyndman, R. J., & Athanasopoulos, G. (2021)

2.2.3. Modelos Mistos

Os modelos mistos, também chamados de modelos híbridos, combinam elementos de modelos explicativos e modelos de séries temporais. Eles incorporam tanto variáveis independentes quanto informações históricas de séries temporais para realizar previsões. Essa abordagem é útil quando se tem um conjunto de dados que

pode ser influenciado por fatores explicativos conhecidos e, ao mesmo tempo, apresenta padrões temporais que precisam ser considerados na previsão. Uma abordagem comum de modelos mistos é o método de regressão com correção de erros (Error Correction Model - ECM), que combina uma equação de regressão com um modelo de correção de erros de séries temporais.

$$DE_{t+1} = f(DE_t, DE_{t-1}, \dots, \text{temperatura}, \text{erro})$$

Fonte: Hyndman, R. J., & Athanasopoulos, G. (2021)

É importante selecionar o modelo mais adequado com base nas características dos dados disponíveis, no conhecimento prévio do fenômeno em estudo e nos objetivos da previsão. Um modelo explicativo é útil porque incorpora informações sobre outras variáveis, porém, há várias razões pelas quais é mais sugerido optar por um modelo de séries temporais em vez de um modelo explicativo ou misto.

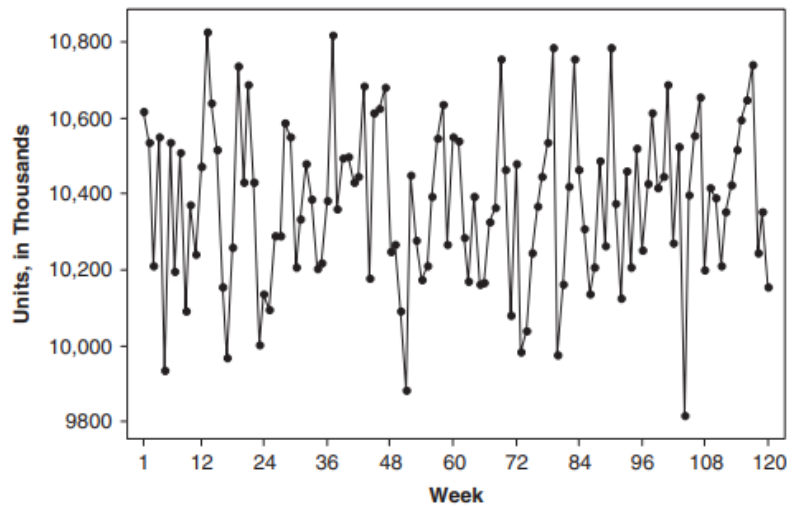
Primeiro, o sistema pode não ser compreendido e, mesmo que fosse, pode ser extremamente difícil medir as relações que supostamente influenciam seu comportamento. Segundo, é necessário conhecer ou prever os valores futuros dos vários preditores para poder prever a variável de interesse, o que pode ser muito difícil. Terceiro, a preocupação principal pode ser apenas prever o que acontecerá, e não saber o porquê acontece. Nesse caso, o modelo de séries temporais pode fornecer previsões mais precisas do que um modelo explicativo ou misto.

2.3. Séries Temporais

2.3.1. Panorama e aplicações

Uma série temporal é definida como "uma coleção de observações ordenadas em intervalos de tempo iguais" (Shumway e Stoffer, 2017). Essas observações podem ser coletadas em diferentes frequências, como minutos, horas, dias, meses ou anos, e podem ser usadas para estudar o comportamento dinâmico e evolutivo de uma variável ao longo do tempo. Abaixo, pode-se observar um exemplo de série temporal:

Gráfico 1 - Série temporal (vendas de uma empresa farmacêutica)



Fonte: Davis, R. A., & Hsu, S. L. (2008)

A análise de séries temporais encontra aplicações em diversas áreas, desempenhando um papel crucial em campos como a previsão de demanda. No contexto empresarial, ela é essencial para o planejamento estratégico de estoques, produção e recursos, proporcionando uma gestão mais eficiente da cadeia de suprimentos (Makridakis, 1998).

Na climatologia e meteorologia, as séries temporais são instrumentos fundamentais. Técnicas especializadas são aplicadas para modelar séries climáticas, incluindo dados como temperatura, precipitação e vento, com o propósito de prever eventos meteorológicos futuros (Chatfield, 2004).

A área da medicina e saúde pública também se beneficia significativamente da análise de séries temporais. Amplamente utilizada, essa abordagem contribui para o monitoramento e previsão da ocorrência de doenças, a identificação de surtos e a avaliação da eficácia de intervenções médicas (Brockwell & Davis, 2016).

2.3.2. Componentes das séries temporais

A previsão de demanda em séries temporais pode ser aprimorada ao considerar quatro componentes principais: tendência, sazonalidade, ruído e ciclos. Esses elementos representam padrões distintos nos dados ao longo do tempo e têm impacto significativo na variação da demanda.

A tendência é um padrão de longo prazo na série temporal que indica uma direção geral de crescimento ou declínio. Ela pode ser ascendente, descendente ou horizontal, influenciando a demanda ao longo do tempo. Métodos de suavização exponencial e modelos de regressão são frequentemente utilizados para identificar e prever tendências (Hyndman & Athanasopoulos, 2018).

Já a sazonalidade refere-se a padrões regulares e previsíveis que se repetem em intervalos de tempo específicos, como semanas, meses ou estações do ano. Essas flutuações periódicas ocorrem em torno de um valor médio e podem ser causadas por fatores sazonais, climáticos, culturais ou eventos específicos (Chatfield, 2004). A detecção e modelagem da sazonalidade são cruciais para prever a demanda em curtos prazos. Métodos como decomposição sazonal e modelos SARIMA são comumente empregados para tratar a sazonalidade.

Ademais, o ruído é um componente que representa a variação aleatória ou o erro residual que não pode ser atribuído às tendências, sazonalidades ou ciclos presentes na série. É a parte não explicada pela modelagem dos outros componentes e pode ser resultado de fatores imprevisíveis ou eventos não controlados que afetam os dados (Box, 2015).

Os componentes clássicos de uma série temporal são: tendência, sazonalidade e ruído. No entanto, alguns autores incluem o "ciclo" como um componente adicional, enquanto outros o incorporam como parte da tendência ou consideram que ele pode ser capturado pelos outros componentes. A presença do ciclo como um componente independente geralmente é mais comum em contextos específicos, como análises econômicas de longo prazo, onde podem ocorrer flutuações cíclicas relacionadas a ciclos econômicos, expansões e recessões.

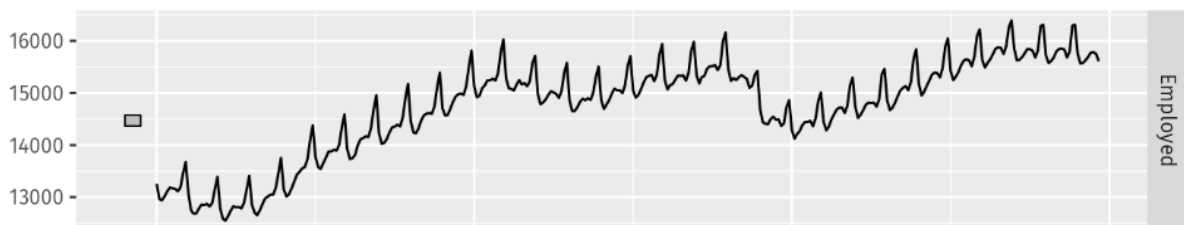
Em contraste, em outras abordagens e aplicativos de séries temporais, os ciclos podem ser considerados como parte da tendência ou como flutuações de longo prazo que são incorporadas no modelo de previsão de forma mais geral. Além disso, em muitos casos, os ciclos podem ser difíceis de identificar e modelar separadamente, especialmente se suas características não forem bem definidas ou ocorrerem em diferentes frequências.

Os ciclos representam flutuações de longo prazo que não são tão regulares quanto a sazonalidade e podem ter durações variáveis. Diferentemente da tendência, os ciclos não seguem uma direção específica de crescimento ou declínio, mas mostram oscilações periódicas em torno de uma linha de tendência. A identificação

de ciclos pode ser complexa e requer técnicas avançadas, como análise de Fourier ou métodos de decomposição de tendência e ciclo (Box, 2015).

Abaixo, podemos observar a decomposição clássica da série temporal. No gráfico 2, temos a série temporal representando o número de empregados no varejo americano durante um certo período. Nos gráficos subsequentes, observamos a influência e o comportamento de cada componente.

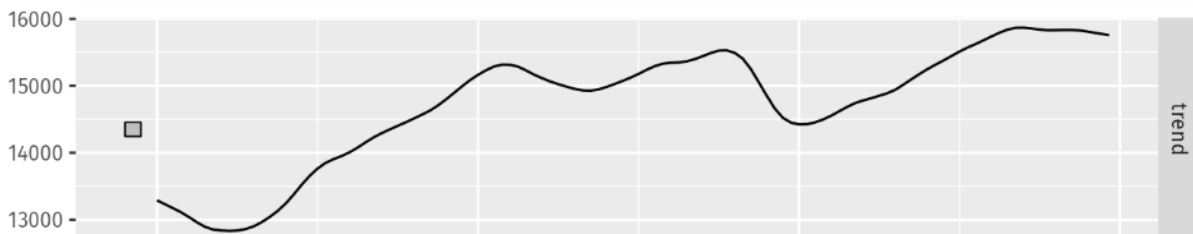
Gráfico 2 – Série temporal (exemplo)



Fonte: Hyndman, R. J., & Athanasopoulos, G. (2021)

O gráfico 2, que representa a série temporal original ao longo do tempo, sem a decomposição em componentes específicos.

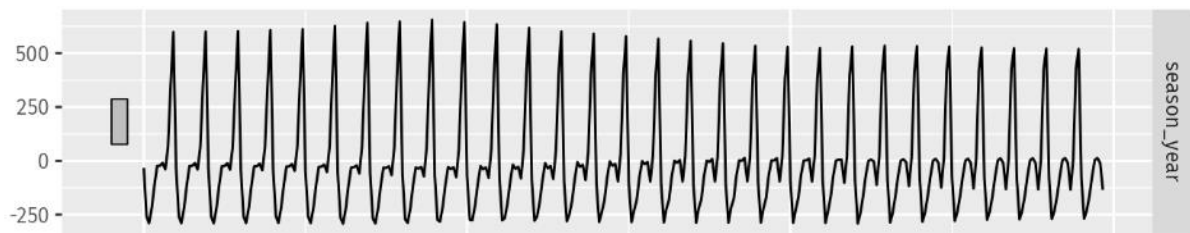
Gráfico 3 – Tendência (exemplo)



Fonte: Hyndman, R. J., & Athanasopoulos, G. (2021)

Já o gráfico 3 representa a tendência da série temporal, que mostra a direção geral ou padrão de comportamento da série ao longo do tempo. Nesse caso, podemos observar tendência de crescimento no número de empregados pelo varejo americano.

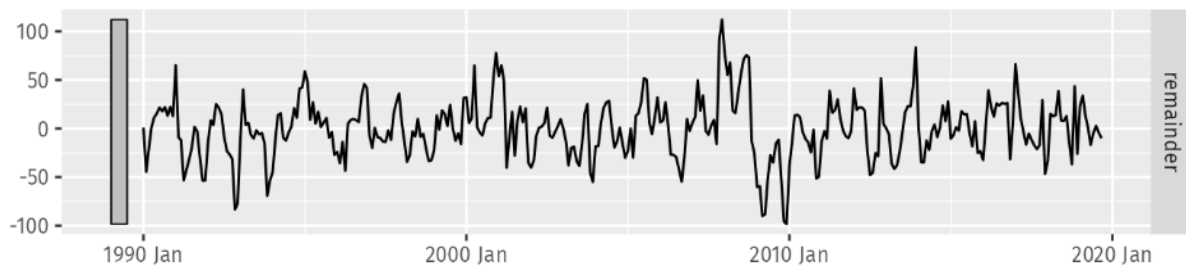
Gráfico 4 – Sazonalidade (exemplo)



Fonte: Hyndman, R. J., & Athanasopoulos, G. (2021)

O gráfico 4 representa a sazonalidade da série temporal, ou seja, padrões que se repetem em intervalos regulares de tempo. Nesse exemplo, observamos um padrão de alta seguido de uma ligeira queda. Após essa queda, existe um crescimento expressivo, acompanhado de uma queda abrupta.

Gráfico 5 – Ruído (exemplo)



Fonte: Hyndman, R. J., & Athanasopoulos, G. (2021)

Por fim, o gráfico 5 representa o ruído da série temporal, ou seja, variações não explicadas pela tendência e sazonalidade. No exemplo acima, os ruídos encontram-se próximos da média 0, o que mostra a captura eficaz dos componentes da série em questão.

2.3.3. Estacionariedade de séries temporais

A estacionariedade é um conceito fundamental na análise de séries temporais e desempenha um papel importante na aplicação de modelos estatísticos, como o ARIMA (AutoRegressive Integrated Moving Average).

Uma série temporal é considerada estacionária quando suas propriedades estatísticas não mudam ao longo do tempo. Isso significa que a média, a variância e a autocorrelação (ou seja, como os valores estão correlacionados com seus atrasos

temporais) permanecem constantes ao longo de todas as observações (Brockwell & Davis, 2016).

A estacionariedade é uma característica crucial em análise de séries temporais, trazendo benefícios significativos. Em séries estacionárias, os padrões e comportamentos ao longo do tempo são mais previsíveis e consistentes, facilitando a aplicação de modelos estatísticos como o ARIMA, os quais pressupõem a estacionariedade dos dados.

Além disso, a interpretação dos parâmetros em modelos ARIMA é mais direta em séries estacionárias, permitindo uma compreensão mais clara do impacto das observações passadas na observação atual.

Existem vários métodos para verificar a estacionariedade de uma série temporal, incluindo análise gráfica - plotar a série temporal e observar se há tendências visíveis, flutuações sazonais ou padrões não aleatórios - e testes estatísticos - testes formais, como o Teste ADF (Augmented Dickey-Fuller) e o Teste KPSS (Kwiatkowski-Phillips-Schmidt-Shin).

Se uma série não for estacionária, é possível torná-la estacionária através de diferenciação. Isso envolve subtrair observações adjacentes ou atrasadas em intervalos regulares. O número de diferenciações necessárias é indicado pelo valor "d" no modelo ARIMA (por exemplo, ARIMA (p, d, q)).

2.3.4. Decomposição de séries temporais

Ao realizar a decomposição dos componentes das séries temporais, é possível obter uma melhor compreensão da série temporal e, conseqüentemente, aprimorar as previsões de demanda. Métodos estatísticos e de aprendizado de máquina são frequentemente empregados para realizar análises de séries temporais e previsões de demanda, considerando os aspectos de tendência, sazonalidade e ciclos (Hyndman & Khandakar, 2008).

Na decomposição por adição, os componentes da série temporal são combinados através de adição. Esse método é geralmente aplicado quando a amplitude das flutuações sazonais é constante ao longo do tempo, sendo adequado para séries temporais em que os efeitos sazonais são mais constantes e não se amplificam com o crescimento ou declínio da série.

$$Y(t) = T(t) + S(t) + R(t) \quad (1)$$

Fonte: Hyndman, R. J., & Athanasopoulos, G. (2021)

Onde:

- $Y(t)$ é o valor observado da série temporal no tempo t .
- $T(t)$ é a tendência, que representa o comportamento de longo prazo da série.
- $S(t)$ é a sazonalidade, que representa os padrões sazonais regulares e repetitivos.
- $R(t)$ é o resíduo ou ruído, que é a variação irregular ou aleatória que não pode ser explicada pelos outros componentes.

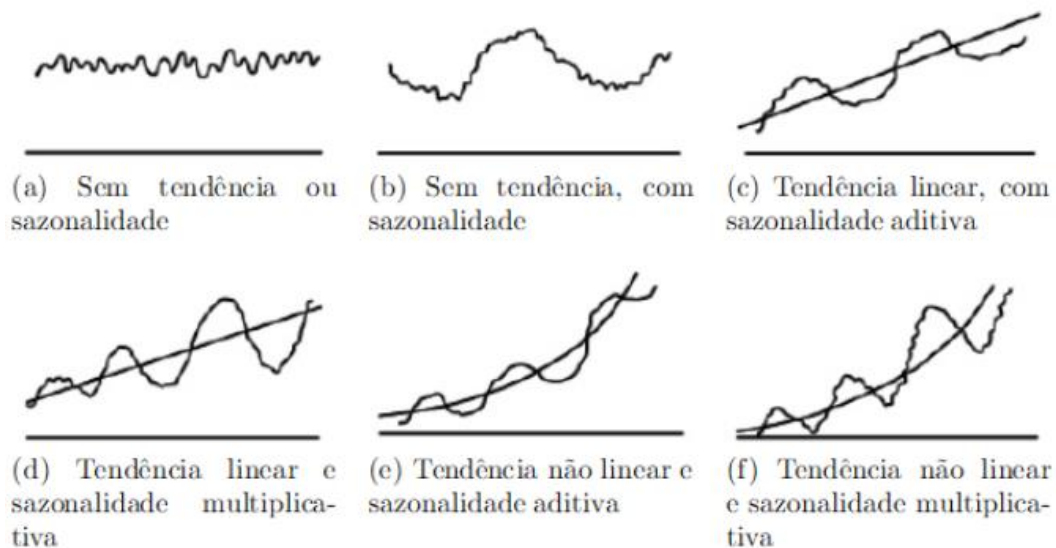
Já na decomposição por multiplicação, os componentes da série temporal são combinados através de multiplicação. Esse método é aplicado quando a amplitude das flutuações sazonais varia ao longo do tempo e é proporcional ao nível da série temporal. A decomposição por multiplicação é adequada para séries temporais em que os efeitos sazonais crescem ou diminuem à medida que o nível da série aumenta ou diminui.

$$Y(t) = T(t) * S(t) * R(t) \quad (2)$$

Fonte: Hyndman, R. J., & Athanasopoulos, G. (2021)

Em alguns casos, a decomposição por multiplicação é preferida porque pode lidar melhor com flutuações sazonais que aumentam ou diminuem com o tempo. Um exemplo de aplicação da decomposição por multiplicação é a decomposição sazonal de X-12-ARIMA, que é uma técnica de ajuste sazonal amplamente utilizada.

Figura 2 - Tendência e sazonalidade para séries temporais



Fonte: Mehl, M. R., Conner, T. S. (2012)

2.4. Autocovariância e autocorrelação

A autocovariância e a autocorrelação desempenham papéis cruciais na análise de séries temporais e são fundamentais para o desenvolvimento de modelos de previsão. Elas oferecem insights valiosos sobre a dependência temporal entre as observações, permitindo a identificação de padrões e comportamentos ao longo do tempo.

2.4.1. Autocovariância

A autocovariância é uma medida estatística que descreve a covariância entre duas observações em uma série temporal, levando em consideração a diferença de tempo entre elas. Ela indica como as observações em diferentes momentos no tempo estão relacionadas em termos de suas flutuações conjuntas (Brockwell & Davis, 2016). A autocovariância entre duas observações X_t e X_{t+k} , onde k é a diferença de tempo (lag), é definida como:

$$Cov(X_t, X_{t+k}) = E[(X_t - \mu)(X_{t+k} - \mu)] \quad (3)$$

Fórmula da autocovariância

Onde:

- X_t e X_{t+k} são as observações nos tempos t e $t+k$, respectivamente.
- μ é a média da série temporal.

A autocovariância é uma medida útil para entender como as flutuações nos dados estão relacionadas ao longo do tempo. Se a autocovariância diminui à medida que a diferença de tempo k aumenta, isso sugere que as observações se tornam menos relacionadas à medida que se afastam umas das outras no tempo.

2.4.2. Autocorrelação

A autocorrelação é uma versão padronizada da autocovariância e é frequentemente usada na análise de séries temporais. Ela mede a relação linear entre as observações em diferentes momentos no tempo, após a remoção do efeito da média e da escala (normalização) (Davis & Hsu, 2018). A função de autocorrelação (ACF) é definida como:

$$ACF(k) = Cov(X_t, X_{t+k}) / Var(X_t) \quad (4)$$

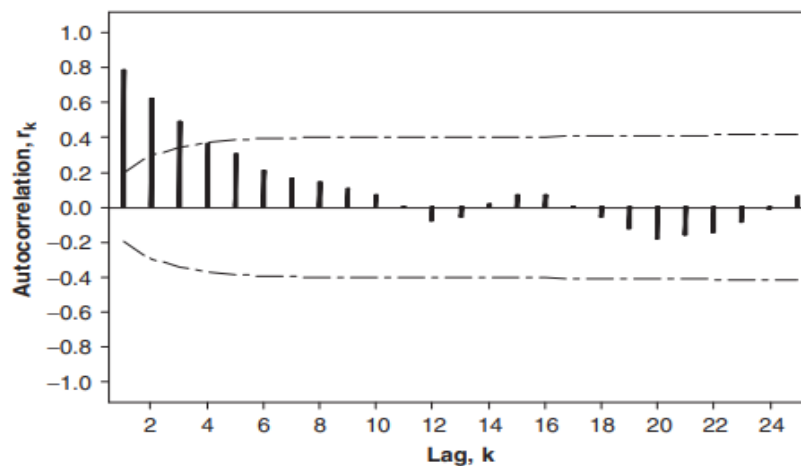
Fórmula da autocorrelação

Onde:

- $ACF(k)$ é a autocorrelação entre as observações X_t e X_{t+k} .
- $Cov(X_t, X_{t+k})$ é a autocovariância entre as observações.
- $Var(X_t)$ é a variância das observações no tempo t .

A função de autocorrelação é frequentemente representada graficamente em um gráfico conhecido como "gráfico de autocorrelação" ou "ACF plot". Ele fornece informações sobre a dependência temporal das observações em diferentes lags (diferenças de tempo):

Gráfico 6 - Autocorrelação para diferentes lags



Fonte: Davis, R. A., & Hsu, S. L. (2008)

Se a autocorrelação em um dado lag k for próxima de 1, isso sugere uma forte correlação positiva entre as observações em t e $t+k$. Se a autocorrelação em um lag k for próxima de -1, isso sugere uma correlação negativa. Se a autocorrelação for próxima de 0, isso indica que as observações são praticamente não correlacionadas.

A principal diferença entre a autocovariância e a autocorrelação está na normalização. A autocovariância mede a covariância bruta entre duas observações em diferentes momentos no tempo, enquanto a autocorrelação normaliza essa covariância, tornando-a independente das escalas das variáveis. A autocorrelação varia de -1 a 1, o que facilita a interpretação da força e da direção da relação linear entre as observações. A autocovariância não possui essa normalização, o que torna mais difícil comparar a força da relação em diferentes momentos no tempo ou entre diferentes séries temporais.

2.4.3. Função de Autocorrelação Parcial

A função de autocorrelação parcial (PACF - Partial Autocorrelation Function) é uma ferramenta importante na análise de séries temporais, especialmente na identificação da ordem do componente autorregressivo (AR) em modelos ARIMA.

A PACF mede a correlação entre uma observação em um determinado ponto no tempo e suas observações passadas, levando em consideração e removendo os efeitos das observações intermediárias. Isso significa que a PACF revela a correlação

direta entre a observação atual e observações passadas, excluindo a influência dos valores intermediários (Box & Jenkins, 2015).

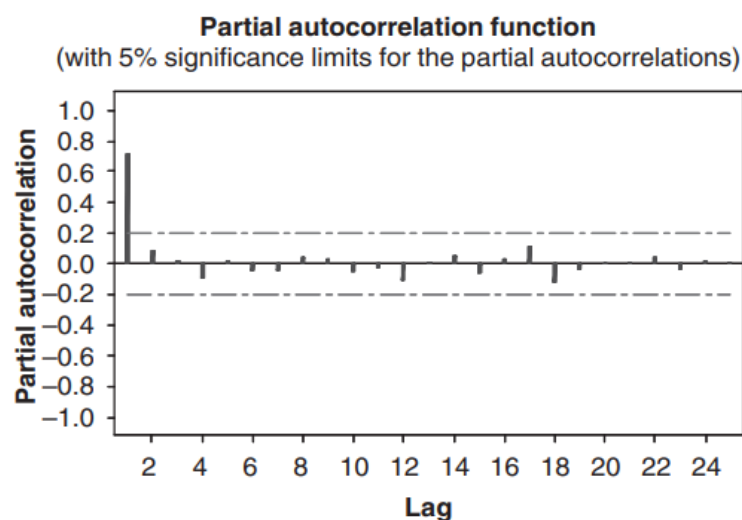
A função de autocorrelação parcial pode ser obtida de maneira iterativa, calculando as autocorrelações parciais para diferentes valores de atraso (lag). A fórmula para a autocorrelação parcial em um determinado atraso $k=3$, por exemplo é dada por:

$$PACF = Cov(X_t, X_{t-3} | X_{t-1}, X_{t-2}) / \sqrt{Var(X_t | X_{t-1}, X_{t-2}) * Var(X_{t-3} | X_{t-1}, X_{t-2})} \quad (5)$$

Fórmula da autocorrelação parcial

Abaixo, temos o gráfico do PACF. Sua interpretação envolve observar os padrões nas autocorrelações parciais em diferentes atrasos (lags):

Gráfico 7 - PACF



Fonte: Davis, R. A., & Hsu, S. L. (2008)

Os lags no gráfico do PACF que se estendem além dos limites de confiança são considerados significativos. Isso significa que a autocorrelação parcial nesses lags é estatisticamente diferente de zero.

A ordem apropriada do componente AR é frequentemente determinada pelo número de lags significativos no gráfico do PACF. Por exemplo, se houver apenas um lag significativo, isso sugere um modelo AR (1).

Em muitos casos, o gráfico do PACF mostra um corte súbito após um determinado lag. Isso sugere que os lags anteriores têm autocorrelações parciais

significativas, mas as autocorrelações parciais dos lags subsequentes são essencialmente zero. Esse corte súbito é uma indicação da ordem do componente AR.

2.5. Modelos de Previsão de Séries Temporais

Nessa seção, serão exploradas uma variedade de abordagens que desempenham um papel crucial na análise preditiva. Começando com a simplicidade da média móvel simples e da regressão linear, avançaremos para técnicas mais sofisticadas, como as estratégias de suavização exponencial simples, dupla e tripla, cada uma adaptada para lidar com diferentes situações de sazonalidade e tendência.

O Modelo ARIMA (AutoRegressive Integrated Moving Average) será introduzido, destacando sua capacidade de lidar com padrões mais complexos, enquanto o Seasonal ARIMA (SARIMA) será explorado para acomodar sazonalidades específicas ao longo do tempo. Essa diversidade de métodos oferecerá uma compreensão abrangente das ferramentas disponíveis para a modelagem e previsão de séries temporais, permitindo uma seleção criteriosa conforme as características específicas de cada conjunto de dados.

2.5.1. Média Móvel Simples

A previsão de demanda por média móvel é uma técnica simples e amplamente utilizada na análise de séries temporais para estimar valores futuros com base na média dos valores passados (Brockwell & Davis, 2016). Essa abordagem assume que o comportamento futuro da série temporal será semelhante ao comportamento médio observado no passado.

A média móvel simples é calculada tomando-se a média aritmética dos últimos n valores da série temporal. A fórmula para o cálculo da média móvel simples em um determinado ponto de tempo t é:

$$SMA = \frac{A_1 + A_2 + \dots + A_n}{n} \quad (6)$$

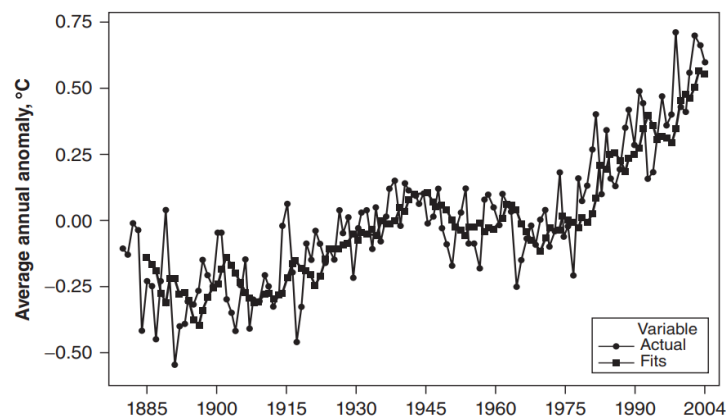
Equação da média móvel

Onde:

- SMA é o valor da média móvel simples no instante t ;
- A_n é a demanda real no período n ;
- n é o número de períodos.

Quando o número de períodos observados é pequeno, a média móvel será mais sensível às variações de curto prazo, pois considerará menos valores passados em sua média. Isso pode levar a uma previsão mais volátil, que seguirá as oscilações da série original de forma mais próxima. A SMA de menor ordem é adequada quando se deseja uma previsão mais responsiva a mudanças recentes na série.

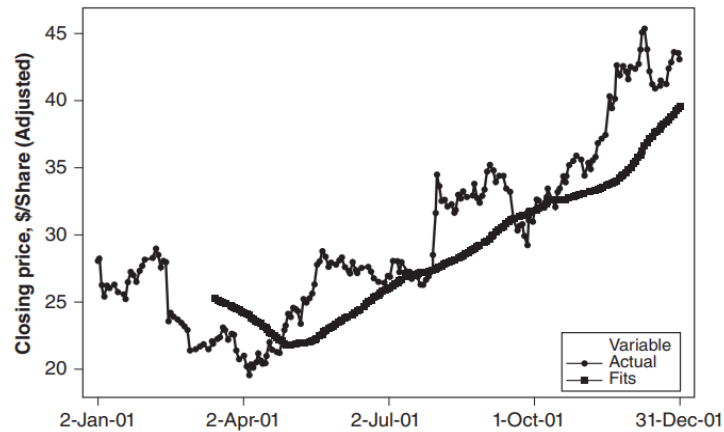
Gráfico 8 - Média móvel de 5 períodos



Fonte: Davis, R. A., & Hsu, S. L. (2008)

Por outro lado, se o número de períodos for maior, a média móvel será mais suavizada e menos sensível a flutuações de curto prazo. A previsão terá menos oscilações e seguirá mais de perto uma tendência ou padrão sazonal de longo prazo. A SMA de maior ordem é útil quando o objetivo é obter uma previsão mais estável e eliminar o ruído da série temporal.

Gráfico 9 - Média móvel de 50 períodos



Fonte: Davis, R. A., & Hsu, S. L. (2008)

Uma limitação das médias móveis é que elas são métodos lineares e não consideram relacionamentos mais complexos entre as variáveis. Em séries temporais com padrões não lineares ou relacionamentos não lineares, as médias móveis podem não ser capazes de capturar adequadamente esses padrões, resultando em previsões imprecisas. Além disso, as médias móveis podem ser sensíveis a outliers presentes na série temporal, afetando a precisão das previsões.

2.5.2. Regressão Linear

A previsão por regressão linear é uma técnica estatística utilizada para estabelecer uma relação linear entre uma variável dependente (ou de resposta) e uma ou mais variáveis independentes (ou de preditores). É frequentemente empregada para fazer previsões ou estimativas com base em dados históricos e nas relações lineares identificadas (Montgomery & Vining, 2012). O modelo de regressão linear é representado pela seguinte equação:

$$y = \beta_0 + \beta_1 * x_1 + \dots + \beta_k * x_k + \varepsilon \quad (7)$$

Equação da regressão linear

Onde:

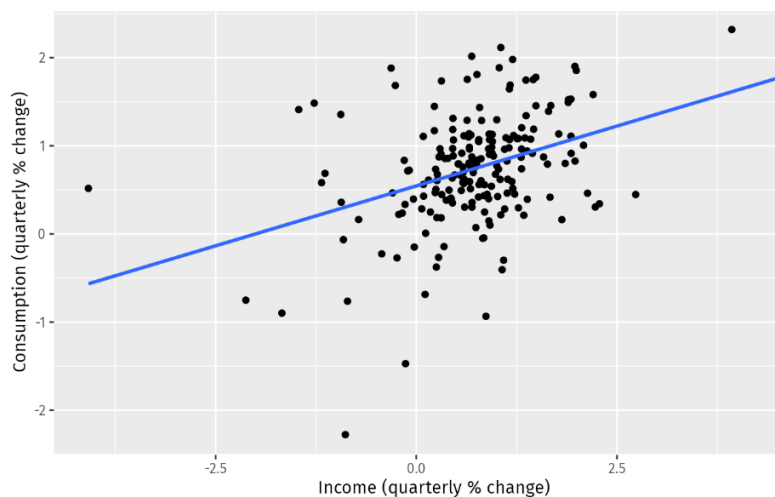
- y é a variável dependente que desejamos prever;
- x_1, x_2, \dots, x_k são as variáveis independentes que podem afetar y ;

- $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ são os coeficientes de regressão que representam as inclinações da reta;
- ε representa o termo de erro, que é a diferença entre os valores observados e os valores previstos pelo modelo.

O objetivo da previsão por regressão linear é estimar os valores dos coeficientes $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ para que o modelo seja capaz de prever y com base nos valores conhecidos de x_1, x_2, \dots, x_k .

Os coeficientes $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ são estimados a partir dos dados históricos através de técnicas de ajuste de regressão, como o método dos mínimos quadrados, que visa minimizar a soma dos quadrados dos resíduos (diferenças entre os valores observados e os valores previstos pelo modelo). Para fazer previsões futuras com o modelo de regressão linear, basta substituir os valores conhecidos de x_1, x_2, \dots, x_k na equação e calcular o valor estimado de y .

Gráfico 10 - Exemplo de regressão linear



Fonte: Hyndman, R. J., & Athanasopoulos, G. (2021)

A previsão por regressão linear pode ser utilizada em diversos contextos, desde previsões de vendas e demanda, até análises financeiras e econômicas. No entanto, é importante ressaltar que a técnica assume uma relação linear entre as variáveis, o que pode ser uma limitação em casos em que as relações são mais complexas.

2.5.3. Suavização Exponencial Simples

A suavização exponencial simples é uma técnica de previsão de séries temporais que calcula previsões futuras com base em uma média ponderada dos valores passados, dando mais peso aos valores mais recentes. Essa abordagem é particularmente útil para séries temporais com padrões de comportamento estáveis e sem padrões sazonais ou tendências significativas (Hyndman & Athanasopoulos, 2021).

A previsão em suavização exponencial simples é baseada em duas etapas: inicialização e atualização. A inicialização envolve a definição de um valor inicial de suavização e uma previsão inicial. A atualização é realizada a cada novo período, utilizando a fórmula de suavização exponencial para obter a previsão para o próximo período. A fórmula para a suavização exponencial simples é representada como:

$$B_t = \alpha * D_t + (1 - \alpha) * B_{t-1} \quad (8)$$

$$F_t(t + k) = B_t$$

$$k = 1, 2, \dots$$

Equação compacta de suavização exponencial simples

Onde:

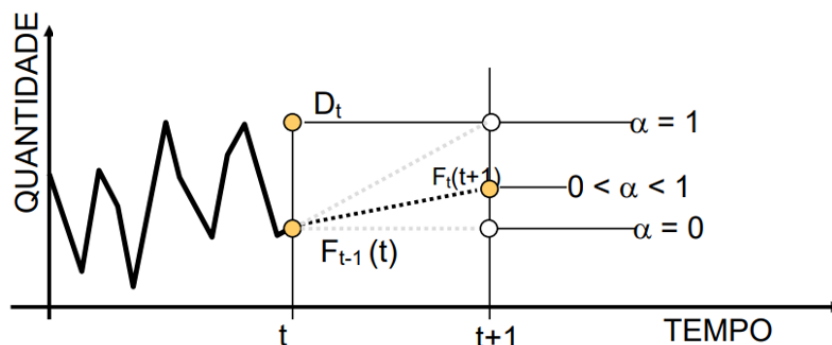
- B_t é a base para previsão no período t ;
- D_t é o valor observado na série temporal no período t ;
- B_{t-1} é a base anterior para previsão, ou seja, a base da previsão para o período $t - 1$;
- α é o fator de suavização, um valor entre 0 e 1 que determina o peso atribuído aos valores passados em relação aos valores mais recentes. Quanto mais próximo de 1, maior o peso dado aos valores recentes;
- F_t é a previsão para o período t .

A escolha adequada do valor de α é fundamental para o desempenho da suavização exponencial simples. Valores altos de α enfatizam mais os valores recentes, tornando a previsão mais responsiva a mudanças recentes nos dados, mas

também mais suscetível a flutuações aleatórias. Valores baixos de α resultam em uma previsão mais suave e estável, mas podem demorar a reagir a mudanças significativas na série.

A suavização exponencial simples é uma técnica fácil de implementar, sendo amplamente utilizada para previsões de curto prazo em séries temporais com comportamento relativamente estável. No entanto, em séries com tendências, sazonalidades ou padrões mais complexos, outras variantes da suavização exponencial, como a suavização exponencial dupla e a suavização exponencial tripla, podem ser mais apropriadas.

Gráfico 11 - Suavização exponencial simples



Fonte: Adaptado de Namhias, S. (2009)

2.5.4. Suavização Exponencial Dupla (Holt)

A suavização exponencial dupla, também conhecida como Método de Holt, é uma técnica de previsão de séries temporais que leva em consideração a tendência dos dados. Ela é uma extensão da suavização exponencial simples, permitindo fazer previsões em séries temporais que apresentam padrões de tendência (Hyndman & Athanasopoulos, 2021). O modelo de suavização exponencial dupla é representado pelas seguintes equações de atualização:

$$B_t = \alpha * D_t + (1 - \alpha) * (B_{t-1} + T_{t-1}) \quad (9)$$

$$T_t = \beta * (B_t - B_{t-1}) + (1 - \beta) * T_{t-1}$$

$$F_t(t + k) = B_t + kT_t$$

$$k = 1, 2, \dots$$

Equações de suavização exponencial dupla

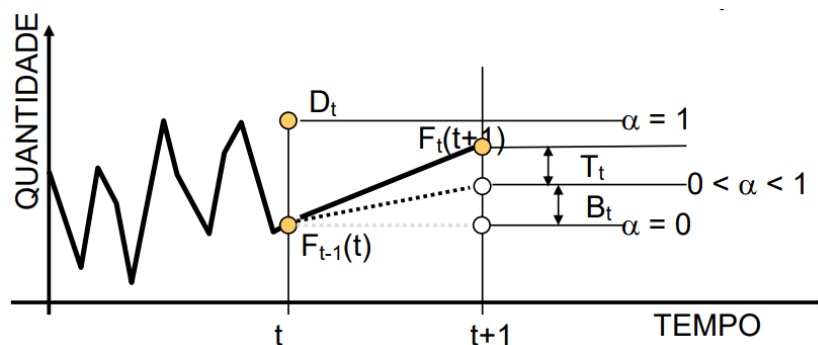
Onde:

- D_t - demanda do período t ;
- B_t - base ao final do instante t ;
- T_t - tendência ao final do instante t ;
- α - constante de suavização para Base;
- β - constante de suavização para Tendência;
- $F_t(u)$ - previsão ao final do período t para o período u ($u > t$).

O método de Holt é uma abordagem iterativa, onde os valores iniciais de Base e Tendência são estimados a partir dos primeiros valores observados da série temporal, e em seguida, as equações de atualização são aplicadas sequencialmente para estimar os valores futuros.

A suavização exponencial dupla é amplamente utilizada em previsão de séries temporais com padrões de tendência moderadamente ou fracamente definidos. No entanto, ela assume que a tendência muda de forma constante ao longo do tempo, o que pode ser uma limitação em séries com comportamentos mais complexos.

Gráfico 12 - Suavização exponencial dupla



Fonte: Adaptado de Namhias, S. (2009)

2.5.5. Suavização Exponencial Tripla (Holt-Winters)

A suavização exponencial tripla, também conhecida como Holt-Winters, é uma técnica de previsão de séries temporais que leva em conta três componentes

principais: tendência, sazonalidade e erro. Essa abordagem é especialmente útil quando a série temporal apresenta padrões sazonais e tendências (Hyndman & Athanasopoulos, 2021). O modelo de suavização exponencial tripla é representado pelas seguintes equações:

$$B_t = \alpha * \frac{D_t}{I_{t-L}} + (1 - \alpha) * (B_{t-1} + T_{t-1}) \quad (10)$$

$$T_t = \beta * (B_t - B_{t-1}) + (1 - \beta) * T_{t-1}$$

$$I_t = \gamma * \frac{D_t}{B_t} + (1 - \gamma) * I_{t-L}$$

$$F_t(t + k) = (B_t + kT_t) * I_{t-L+k}$$

$$k = 1, 2, \dots, L$$

Equações de suavização exponencial tripla

Onde:

- D_t - demanda do período t;
- B_t - base ao final do instante t;
- T_t - tendência ao final do instante t;
- I_t - índice de sazonalidade do instante t;
- α - constante de suavização para Base;
- β - constante de suavização para Tendência;
- γ - constante de suavização para Sazonalidade;
- $F_t(u)$ - previsão ao final do período t para o período u ($u > t$);
- L - tempo de uma estação completa.

As equações de nível, tendência e sazonalidade são usadas para atualizar os valores dos componentes em cada período. A previsão é feita através da combinação desses componentes, projetando-os k períodos no futuro. O método de suavização exponencial tripla é adequado para séries temporais com tendências e padrões sazonais, e pode ser aplicado em diferentes contextos, como previsões de vendas, demanda, produção e outras séries temporais com comportamentos sazonais e tendências.

2.5.6. Modelo ARIMA

De acordo com Box e Jenkins (2015), o Modelo ARIMA (AutoRegressive Integrated Moving Average) é uma abordagem sofisticada que combina três componentes essenciais para modelagem de séries temporais: o componente autorregressivo (AR), a média móvel (MA) e a diferenciação (I). A equação do ARIMA (p,d,q) é expressa por:

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (11)$$

Equação do modelo ARIMA

Onde:

- p - ordem do componente autorregressivo (AR), que representa o número de lags incluídos no modelo;
- d - ordem de diferenciação, que representa o número de vezes que a série temporal foi diferenciada para torná-la estacionária;
- q - ordem do componente de média móvel (MA), que representa o número de lags dos erros incluídos no modelo;
- Y_t - é o valor da série temporal no tempo t ;
- ϕ_1, \dots, ϕ_p são os coeficientes autorregressivos;
- $\theta_1, \dots, \theta_q$ são os coeficientes da média móvel;
- ε_t - termo de erro no tempo t .

O componente AR mede como a observação atual da série temporal está relacionada às observações anteriores em termos de atrasos (ou lags). A ordem do componente AR é representada como "p" e indica quantos lags passados são considerados no modelo. Por exemplo, um modelo AR (2) inclui as duas observações imediatamente anteriores. Existem alguns passos para identificar o parâmetro "p" mais adequado do modelo.

Primeiro, começar traçando o gráfico da série temporal original e observar se há algum padrão de comportamento que sugira uma relação com observações passadas. Depois, examinar o gráfico ACF para identificar lags (atrasos) com autocorrelações significativas. Lags com autocorrelações que ultrapassam as faixas

de confiança (geralmente 95% ou 99%) podem ser candidatos a "p" no componente AR. Por fim, usar o gráfico PACF para identificar lags com autocorrelações parciais significativas. O primeiro pico no PACF que não é explicado pelo pico anterior pode indicar a ordem "p" apropriada para o componente AR.

Já o componente MA modela como a observação atual depende dos erros (ou resíduos) passados em termos de atrasos (lags). A ordem do componente MA é representada como "q" e indica quantos erros passados são considerados no modelo. Um modelo MA (1) considera apenas o erro mais recente. Existem alguns passos para identificar o parâmetro "q" mais adequado do modelo.

Após identificar a ordem "p" para o componente AR, repetir o processo para o componente MA, traçando o gráfico da série temporal original, o que pode fornecer insights sobre padrões que podem estar relacionados aos erros passados. Depois, examinar o gráfico ACF novamente, mas agora, concentrando nos lags com autocorrelações significativas. Lags com autocorrelações fora das faixas de confiança podem ser candidatos a "q" no componente MA. Após isso, usar o gráfico PACF para identificar lags com autocorrelações parciais significativas. O primeiro pico no PACF que não é explicado pelo pico anterior pode indicar a ordem "q" adequada para o componente MA.

Para determinar se será utilizado um modelo AR ("p" maior que 0), MA ("q" maior que 0) ou ARMA ("p" e "q" maiores que 0), é possível basear-se no quadro a seguir, fornecido por Montgomery, Jennings e Kulahci no livro "Introduction to Time Series Analysis and Forecasting":

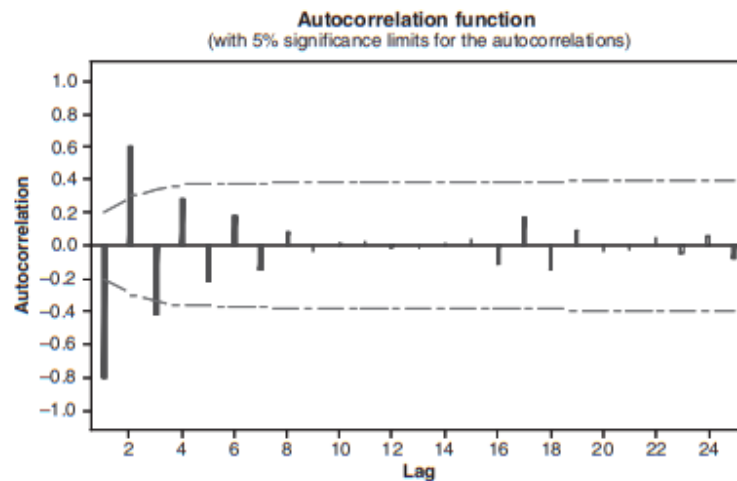
Quadro 1 - Comportamento teórico do ACF e PACF para séries estacionárias

Modelo	ACF	PACF
AR(p)	Decaimento exponencial e/ou senoide amortecida	Corte súbito após lag p
MA(q)	Corte súbito após lag q	Decaimento exponencial e/ou senoide amortecida
ARMA(p, q)	Decaimento exponencial e/ou senoide amortecida	Decaimento exponencial e/ou senoide amortecida

Fonte: Montgomery, Jennings & Kulahci (2015)

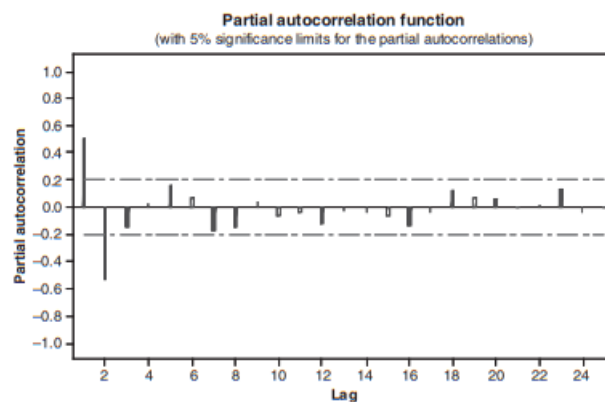
Abaixo, podemos observar o comportamento dos gráficos de ACF e PACF para um modelo AR(2), visto que o ACF possui comportamento senoidal amortecido e o PACF possui um corte súbito.

Gráfico 13 - ACF para um modelo AR(2)



Fonte: Montgomery, Jennings & Kulahci (2015)

Gráfico 14 - PACF para um modelo AR(2)



Fonte: Montgomery, Jennings & Kulahci (2015)

Por fim, a diferenciação refere-se ao processo de tornar a série temporal estacionária, o que é necessário para aplicar o ARIMA. A ordem de diferenciação é representada como "d" e indica quantas vezes a série é diferenciada para atingir a estacionariedade. Isso geralmente envolve calcular a diferença entre uma observação e a observação anterior. Existem alguns passos para identificar o parâmetro "d" mais adequado do modelo.

Antes de determinar a ordem da diferenciação, testar a estacionariedade da série temporal usando análise visual ou o teste Augmented Dickey-Fuller (ADF), por exemplo. Se a série não for estacionária, será necessária a diferenciação. Depois, aplicar diferenciação uma vez ($d = 1$) e testar novamente a estacionariedade.

Continuar aplicando diferenciação até que a série seja estacionária. A ordem "d" é o número de vezes que a diferenciação é necessária para tornar a série estacionária.

Para selecionar os parâmetros de maior precisão do modelo, é possível utilizar algumas métricas, como o AIC, o BIC ou o RMSE. O Critério de Informação Akaike (AIC) é uma métrica usada para comparar modelos estatísticos, incluindo modelos ARIMA, com o objetivo de escolher o modelo que melhor se ajusta aos dados e, ao mesmo tempo, evita a complexidade excessiva. Quanto menor o valor do AIC, melhor o modelo é considerado em termos de equilíbrio entre ajuste e complexidade.

O Critério de Informação Bayesiano (BIC), também conhecido como Critério de Schwarz, é semelhante ao AIC, mas aplica uma penalização maior à complexidade do modelo. A ideia é evitar o overfitting (ajuste excessivo) mais rigorosamente do que o AIC. Já a Raiz do Erro Médio Quadrático (RMSE) calcula a média das diferenças entre as previsões do modelo e os valores reais dos dados. Quanto menor o RMSE, melhor o modelo é em termos de precisão.

Em resumo, o AIC e o BIC são critérios de seleção de modelo que buscam equilibrar o ajuste aos dados e a complexidade do modelo, com o BIC sendo mais rigoroso em relação à complexidade. O RMSE é uma métrica de erro que avalia a precisão das previsões do modelo em relação aos dados reais.

Adicionalmente, é possível empregar técnicas mais avançadas, como o método de "Grid Search" aliado à Validação Cruzada, a fim de identificar o conjunto ideal de parâmetros. Nessa abordagem, realiza-se uma exploração sistemática por meio de uma grade de diferentes combinações de parâmetros para o modelo ARIMA. Para cada conjunto de parâmetros, o modelo é ajustado aos dados e, em seguida, avalia-se o desempenho utilizando técnicas de validação cruzada. Esse processo de validação cruzada proporciona uma avaliação mais robusta do modelo, considerando diferentes divisões dos dados em conjuntos de treinamento e teste, contribuindo para uma seleção mais precisa dos parâmetros que resultam em melhor desempenho preditivo.

2.5.7. Modelo SARIMA

De acordo com Box e Jenkins (2015), o Modelo SARIMA (Seasonal AutoRegressive Integrated Moving Average) é uma extensão do ARIMA que leva em consideração a sazonalidade presente em muitas séries temporais. Além dos

componentes AR, I e MA do ARIMA, o SARIMA também inclui três componentes sazonais adicionais: o componente Sazonal Autorregressivo (SAR), o componente Sazonal de Diferenciação (D), e o componente Sazonal de Média Móvel (SMA). A equação do SARIMA (p,d,q)(P,D,Q,s) pode ser expressa por:

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t + \phi_1 Y_{t-s} + \dots + \phi_P Y_{t-Ps} + \Theta_1 \varepsilon_{t-s} + \dots + \Theta_Q \varepsilon_{t-Qs} \quad (12)$$

Equação do modelo SARIMA

Onde:

- p - ordem do componente autorregressivo (AR) para os lags não sazonais;
- d - ordem de diferenciação não sazonal;
- q - ordem do componente de média móvel (MA), para os lags dos erros não sazonais;
- P - ordem do componente autorregressivo (AR) para os lags sazonais;
- D - ordem de diferenciação sazonal;
- Q - ordem do componente de média móvel (MA), para os lags dos erros sazonais;
- s - período sazonal, que representa o número de observações em uma temporada completa;
- Y_t - é o valor da série temporal no tempo t ;
- ϕ_1, \dots, ϕ_p são os coeficientes autorregressivos não sazonais;
- $\theta_1, \dots, \theta_p$ são os coeficientes de média móvel não sazonal;
- Φ_1, \dots, Φ_P são os coeficientes autorregressivos sazonais;
- $\Theta_1, \dots, \Theta_P$ são os coeficientes da média móvel sazonal;
- ε_t - termo de erro no tempo t .

O componente Sazonal Autorregressivo (SAR) é similar ao componente AR, entretanto, captura a relação linear entre uma observação e suas observações passadas em lags sazonais. A ordem do componente SAR é representada por "P". Para identificar o parâmetro "P" mais adequado do modelo, os seguintes passos podem ser seguidos.

Inicialmente, é recomendável traçar o gráfico da série temporal original e verificar a presença de padrões sazonais evidentes, como picos regulares em determinados intervalos de tempo. Em seguida, o gráfico de Função de Autocorrelação Sazonal (SACF) é utilizado para identificar lags sazonais com autocorrelações significativas. Lags cujas autocorrelações ultrapassem as faixas de confiança sazonal podem ser considerados candidatos para o parâmetro "P" no componente SAR.

O componente Sazonal de Diferenciação (I) é análogo à diferenciação não sazonal, mas é aplicado em lags sazonais. A ordem do componente de diferenciação sazonal é indicada por "D". A identificação do parâmetro "D" mais adequado do modelo segue alguns passos.

Antes de determinar a ordem da diferenciação sazonal, é importante realizar testes de estacionariedade nos dados sazonais usando o teste Augmented Dickey-Fuller (ADF) ou outro teste apropriado. Se os dados sazonais não forem estacionários, é necessária a aplicação da diferenciação sazonal. Posteriormente, a diferenciação sazonal é aplicada uma vez ($D = 1$), e é novamente testada a estacionariedade sazonal. O processo é repetido até que os dados se tornem estacionários em termos sazonais, sendo que a ordem "D" corresponde ao número de vezes que a diferenciação sazonal é necessária para atingir a estacionariedade.

Por fim, o componente Sazonal de Média Móvel (SMA) é similar ao componente MA, porém, captura a relação linear entre uma observação e os erros de previsão em lags sazonais. A ordem do componente SMA é denotada por "Q". Para identificar o parâmetro "Q" mais adequado do modelo, os seguintes passos são seguidos.

Após identificar a ordem "P" para o componente SAR, o processo é repetido para o componente SMA. Neste caso, é recomendável examinar o gráfico da série temporal original em busca de padrões relacionados a erros sazonais. Em seguida, o gráfico SACF é utilizado novamente, desta vez focando nos lags com autocorrelações sazonais significativas. Lags cujas autocorrelações ultrapassem as faixas de confiança sazonal podem ser considerados candidatos para o parâmetro "Q" no componente SMA.

Assim como para o modelo ARIMA, também é possível empregar o método de "Grid Search" aliado à Validação Cruzada, para identificar o conjunto ideal de parâmetros. Nessa abordagem, realiza-se uma exploração sistemática por meio de

uma grade de diferentes combinações de parâmetros para o modelo SARIMA. Para cada conjunto de parâmetros, o modelo é ajustado aos dados e, em seguida, avalia-se o desempenho utilizando técnicas de validação cruzada.

2.6. Erros nos Métodos de Previsão

Na avaliação de modelos de previsão de séries temporais, a análise dos erros desempenha um papel fundamental para entender o desempenho e a confiabilidade das projeções. Diversas métricas são empregadas para mensurar a discrepância entre as previsões e os valores reais, oferecendo insights valiosos sobre a precisão do modelo.

Nesta seção, examinaremos quatro métricas: Erro Médio (ME), Erro Absoluto Médio (MAE), Erro Percentual Absoluto Médio (MAPE) e a Raiz do Erro Quadrático Médio (RMSE). Cada uma dessas métricas fornece uma perspectiva única sobre diferentes aspectos do desempenho do modelo, permitindo uma avaliação abrangente das capacidades de assertividade na previsão.

2.6.1. Erro Médio (ME)

O erro médio, "Mean Error" em inglês, é uma métrica de avaliação usada para medir a precisão de um modelo de previsão ou estimativa em relação aos valores reais. Ele é calculado como a média dos desvios entre as previsões feitas pelo modelo e os valores reais. A fórmula geral para calcular o erro médio é a seguinte:

$$ME = \sum_{t=1}^N \frac{(D_t - F_t)}{N} \quad (13)$$

Equação do Erro Médio

Onde:

- D_t : Os valores reais ou observados na amostra.
- F_t : Os valores previstos pelo modelo.
- N : O número total de observações na amostra.

O erro médio pode ser positivo ou negativo, dependendo se o modelo tende a superestimar ou subestimar os valores reais. Um erro médio próximo de zero indica que o modelo está fazendo previsões que, em média, são próximas aos valores reais.

É importante notar que o erro médio não leva em consideração a magnitude dos desvios individuais; ele apenas calcula a média dos desvios. Desse modo, ele pode não capturar completamente a qualidade geral de um modelo de previsão, especialmente se houver desvios substanciais em algumas previsões.

2.6.2. Erro Absoluto Médio (MAE)

O Erro Absoluto Médio (MAE, em inglês) mede a magnitude média das diferenças absolutas entre os valores previstos e os valores reais. Em outras palavras, o MAE calcula a média das diferenças positivas (sem considerar a direção) entre as previsões e os dados reais. A fórmula para calcular o Erro Absoluto Médio é a seguinte:

$$MAE = \sum_{t=1}^N \frac{|D_t - F_t|}{N} \quad (14)$$

Equação do Erro Absoluto Médio

Onde:

- D_t : Os valores reais ou observados na amostra.
- F_t : Os valores previstos pelo modelo.
- N : O número total de observações na amostra.

Quanto menor o MAE, mais precisa é a previsão, pois indica que as previsões estão mais próximas dos valores reais. Por outro lado, um MAE maior indica que as previsões têm uma maior dispersão em relação aos valores reais.

2.6.3. Erro Percentual Absoluto Médio (MAPE)

O Erro Percentual Absoluto Médio (MAPE, em inglês) mede a precisão de um modelo de previsão em relação aos valores reais e expressa o erro como uma porcentagem da demanda real. O MAPE é especialmente útil quando se deseja

entender o quão precisas são as previsões em termos percentuais, o que facilita a interpretação do erro. A fórmula para calcular o MAPE é a seguinte:

$$MAPE = \sum_{t=1}^N \frac{|D_t - F_t|}{D_t} \quad (15)$$

Equação do Erro Percentual Absoluto Médio

Onde:

- D_t : Os valores reais ou observados na amostra.
- F_t : Os valores previstos pelo modelo.
- N : O número total de observações na amostra.

Quanto menor o valor do MAPE, mais preciso é o modelo de previsão em termos percentuais. No entanto, é importante observar que essa métrica pode não ser a métrica ideal em todos os cenários, especialmente quando os valores reais são muito próximos de zero, pois isso pode resultar em valores infinitos ou indefinidos no denominador.

2.6.4. Raiz do Erro Quadrático Médio (RMSE):

A Raiz do Erro Quadrático Médio (RMSE, em inglês) é uma métrica de avaliação que representa a raiz quadrada da média dos quadrados dos desvios entre os valores previstos pelo modelo e os valores reais. O RMSE é uma medida de como os erros de previsão se distribuem em torno da média e é especialmente útil para penalizar erros maiores de forma mais intensa. A fórmula do RMSE é a seguinte:

$$RMSE = \sqrt{\frac{\sum_{t=1}^N (D_t - F_t)^2}{N}} \quad (16)$$

Equação da Raiz do Erro Quadrático Médio

Onde:

- D_t : Os valores reais ou observados na amostra.
- F_t : Os valores previstos pelo modelo.

- N : O número total de observações na amostra.

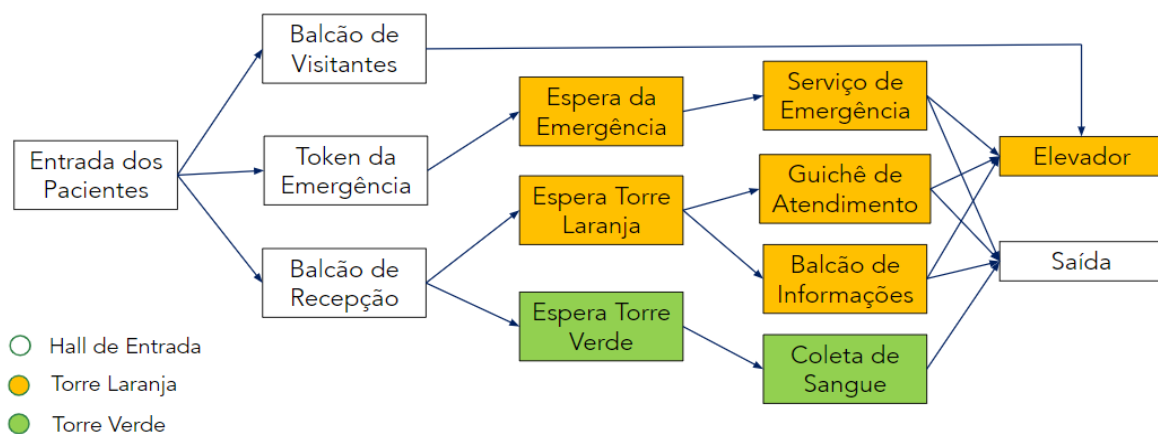
Valores menores de RMSE indicam que o modelo tem um desempenho melhor, pois os erros de previsão são menores e mais próximos dos valores reais. No entanto, uma desvantagem dessa métrica é que ele é sensível a valores extremos, ou seja, erros muito grandes podem distorcer significativamente o resultado. Portanto, ao interpretar o RMSE, é importante considerar a escala dos dados e o contexto do problema. Como o RMSE é expresso na mesma unidade de medida que os valores reais, facilita a interpretação do erro médio em termos da unidade original da variável de interesse.

3. ANÁLISE DA CHEGADA DOS PACIENTES

3.1. Descrição do Processo

Para entender o processo de chegadas, foram realizadas algumas visitas de campo no ICESP, conversando com os administradores do hospital, atendentes e pacientes. O fluxograma abaixo auxilia no entendimento de como funcionam os atendimentos no andar térreo do instituto:

Figura 3 - Fluxograma dos atendimentos no térreo

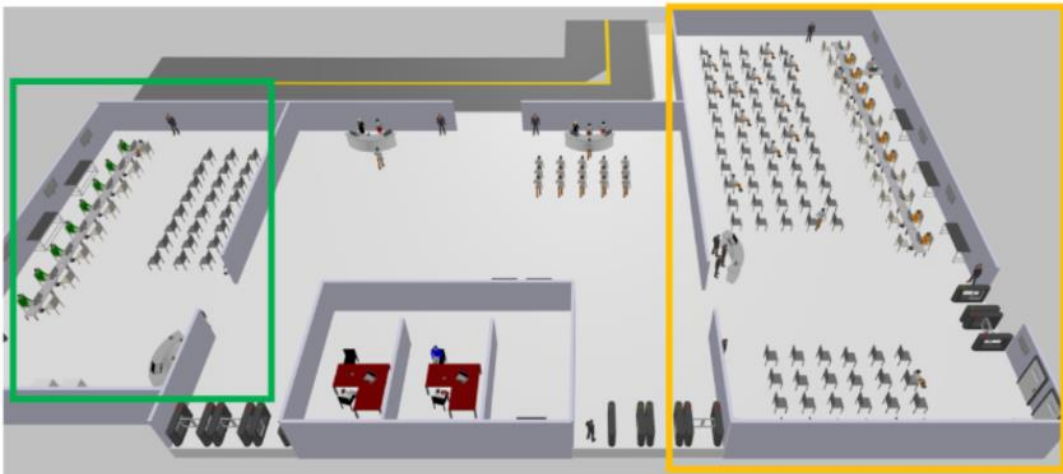


Fonte: formulação própria

Os pacientes entram no hospital com 3 possibilidades de atendimento inicial. A primeira delas é o balcão de visitantes no hall de entrada, no qual é feito o cadastro do visitante, que em seguida vai ao elevador para realizar a visita ao leito desejado. Outra possibilidade é a ida até o token de emergência, que fica dentro da torre laranja - focada nos procedimentos como consultas, exames e quimioterapia.

Por fim, a última possibilidade - também a mais comum delas - é a ida ao balcão da recepção. Após esse primeiro atendimento no balcão, o paciente pode ir à torre verde - responsável pela coleta de sangue - ou para a torre laranja, a fim de realizar o procedimento desejado. Em seguida, o paciente vai à torre designada e realiza seu atendimento já no térreo (como no caso da coleta de sangue) ou é encaminhado ao elevador.

Figura 4 - Representação 3D do fluxo de atendimento no térreo



Fonte: formulação própria (software de simulação Anylogic)

Figura 5 - Guichês de atendimento na torre laranja



Fonte: Youtube

3.2. Análise dos Dados

Para realizar o estudo sobre a análise da chegada dos pacientes do hospital, tivemos acesso a uma base de dados fornecida pelo ICESP, com informações de abril de 2019 a agosto de 2019. Primeiramente, foi realizado um comparativo entre os possíveis destinos que o paciente pode ter ao chegar no hospital. De acordo com a tabela abaixo, 71.5% dos pacientes têm como destino a torre laranja, que é responsável pelos atendimentos e consultas no geral.

O fluxo de maior representatividade no processo é a recepção da torre laranja (57.6% do total de chegadas), que realiza o primeiro atendimento dos pacientes, que chegam nos guichês e são encaminhados aos andares superiores para o procedimento ou consulta com o médico.

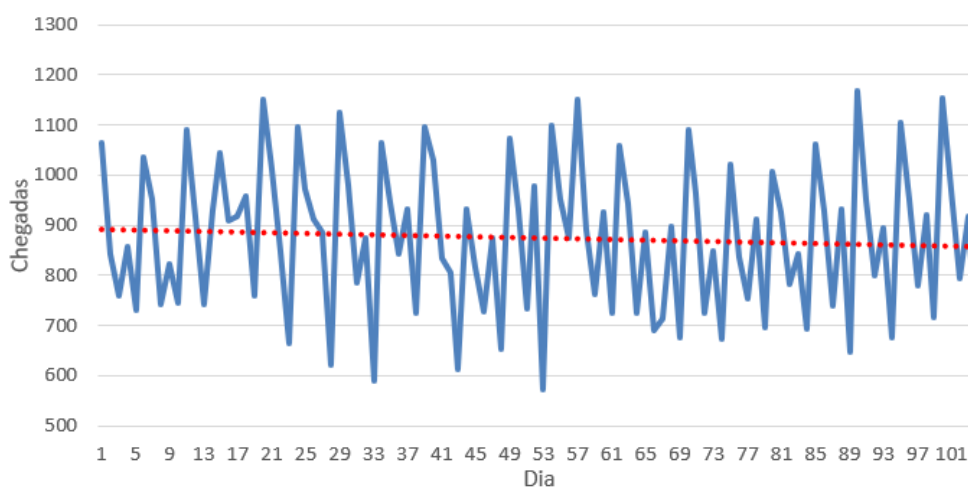
Quadro 2 - Distribuição de destinos

Destino	Torre	Distribuição	
Aval. Doc. Informações Gerais	Laranja	7.1%	71.5%
Convocação para Exames		1.0%	
Pré - Atendimento		5.9%	
Recepção Torre Laranja		57.6%	
Retirada de coletor	Verde	0.7%	28.5%
Cadastro Coleta de Sangue		27.8%	

Fonte: elaborado pelo autor

Visto que as chegadas na recepção da torre laranja são as mais significativas - além de representarem um ponto de preocupação para os administradores do ICESP devido à formação de longas filas - o foco de estudo será em relação a esse destino. Abaixo, é possível observar o padrão de chegadas na recepção da torre laranja. Para facilitar a análise, removemos os dias com chegadas nulas, já que o ICESP não funciona aos finais de semana (resultando em 104 observações diárias).

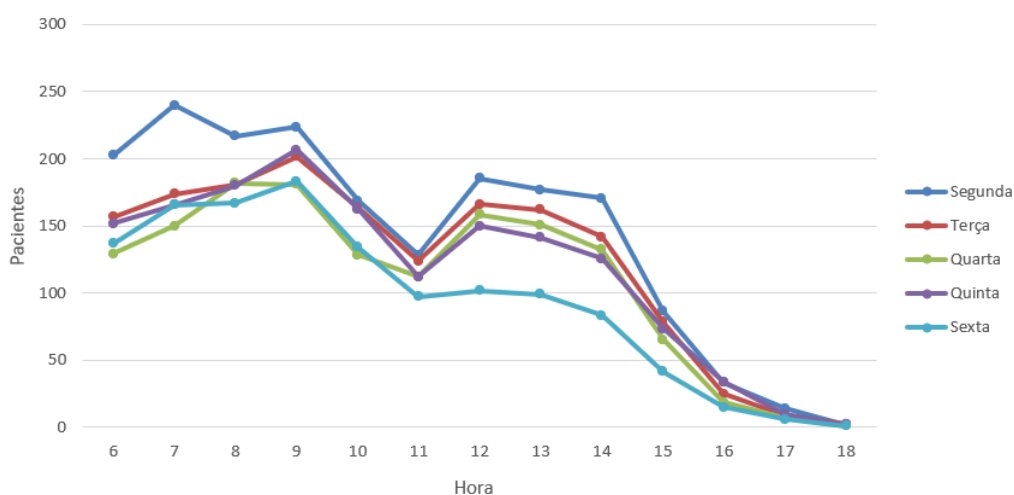
Gráfico 15 - Chegadas por dia (Recepção Torre Laranja)



Fonte: elaborado pelo autor

O comportamento das chegadas por dia oscila entre 600 e 1200 chegadas, em torno da média de 900 chegadas. O dia da semana com mais chegadas é segunda-feira, tendo uma tendência de queda nos atendimentos até sexta-feira. Além disso, as horas do dia com mais chegadas são de 6 a 9 horas, tendo uma diminuição até meio-dia. Após o meio-dia, existe um aumento nas chegadas, que se mantém elevadas até as 15 horas, com uma tendência de queda após esse período.

Gráfico 16 - Média de chegadas por dia da semana e por hora



Fonte: elaborado pelo autor

Quadro 3 - Média de chegadas por dia da semana e por hora

Hora	Segunda	Terça	Quarta	Quinta	Sexta
6	204	158	130	153	138
7	240	174	150	166	167
8	217	182	182	180	167
9	224	201	181	207	182
10	168	163	128	162	134
11	127	124	112	112	96
12	186	165	158	150	102
13	177	162	151	141	99
14	170	142	132	126	83
15	86	78	65	73	41
16	33	25	18	34	15
17	14	9	7	10	6
18	2	2	1	2	1
Total	1848	1585	1417	1514	1232

Fonte: elaborado pelo autor

Ademais, o comportamento das chegadas é similar quando a comparação é feita por mês - para todos os meses observados na análise, as horas do dia com mais chegadas são de 6 a 9 horas, tendo uma diminuição até meio-dia, voltando a crescer a partir das 12 horas, com tendência de queda a partir das 14 horas.

Gráfico 17 - Média de chegadas por mês e por hora



Fonte: elaborado pelo autor

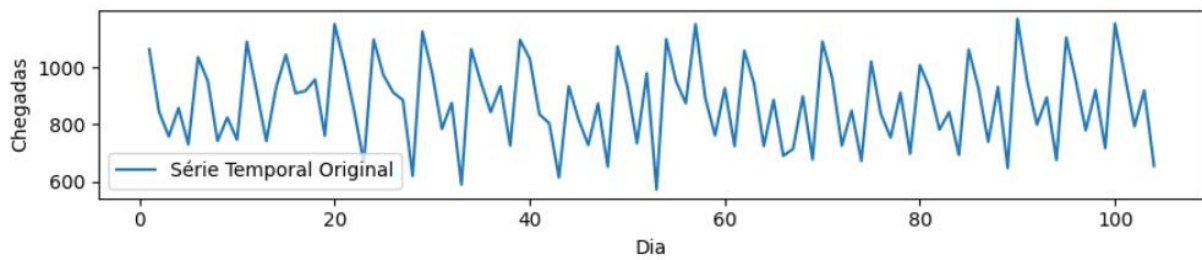
Quadro 4 - Média de chegadas por mês e por hora

Hora	Abril	Maio	Junho	Julho	Agosto
6	151	147	154	164	167
7	177	173	180	181	186
8	182	183	181	186	197
9	196	192	198	197	213
10	140	143	156	148	167
11	121	117	120	110	106
12	165	156	152	145	143
13	155	159	148	135	133
14	147	133	133	122	118
15	71	68	73	65	67
16	25	24	26	25	26
17	9	9	9	11	9
18	1	2	1	2	2
Total	1,540	1,505	1,531	1,490	1,531

Fonte: elaborado pelo autor

Para compreender melhor o padrão de chegadas dos pacientes, foi realizada a decomposição da série temporal, com o auxílio da biblioteca “statsmodels” do Python.

Gráfico 18 - Série temporal original

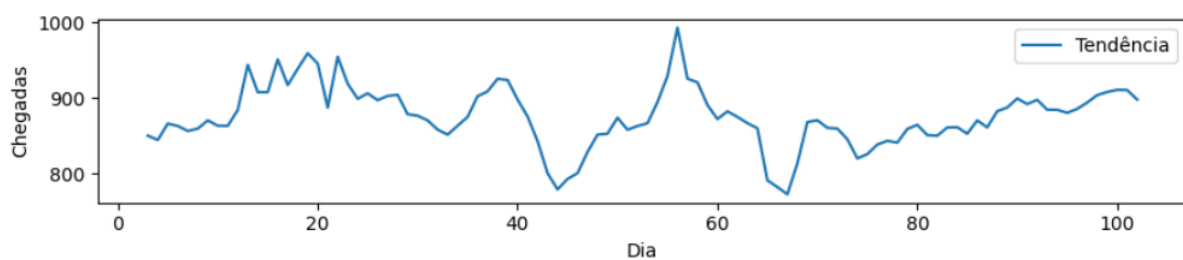


Fonte: elaborado pelo autor

Com o auxílio da mesma biblioteca do Python, foi realizado o teste de estacionariedade de Dickey-Fuller, que produziu um p-valor de 1.8%. Portanto, a série é considerada estacionária (por estar abaixo do limite de confiança de 5%). Além disso, podemos confirmar visualmente que a série é estacionária pois sua média permanece aproximadamente constante ao longo do tempo, sem tendência óbvia de aumento ou diminuição.

A estacionariedade de uma série temporal é fundamental pois facilita a modelagem, torna as previsões mais precisas e reduz o ruído nos dados. Em séries estacionárias, os padrões e propriedades estatísticas são mais consistentes ao longo do tempo. Isso é particularmente útil em modelos estatísticos, como o ARIMA, que pressupõem estacionariedade.

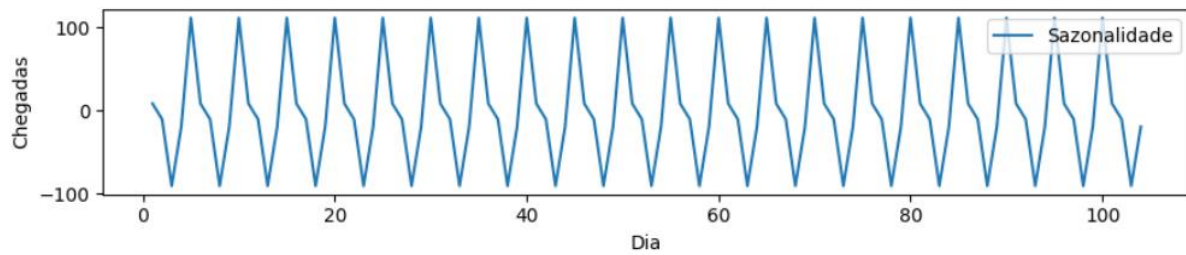
Gráfico 19 - Tendência



Fonte: elaborado pelo autor

O gráfico de tendência apresenta mudanças mais significativas entre os dias 40 e 70, com quedas e subidas bruscas (saindo de 800 até atingir 1000 chegadas diárias). Após o dia 70, a tendência da série temporal volta a apresentar comportamento mais previsível, sem alterações significativas, em torno de uma média de 900 chegadas diárias.

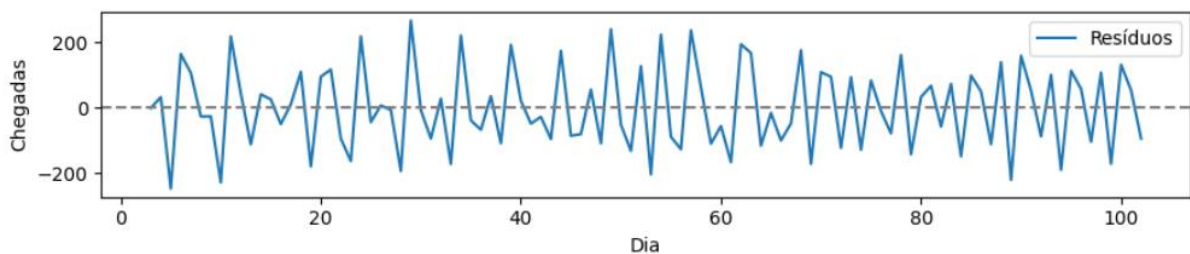
Gráfico 20 - Sazonalidade



Fonte: elaborado pelo autor

Já o gráfico de sazonalidade apresenta padrão bem definido, que pode ser interpretado como a influência do dia da semana no nível de chegadas. Como dito anteriormente, o início da semana (segunda-feira) concentra a maior taxa de chegadas diárias (1848, em média). No decorrer dos dias, a taxa de chegadas decresce, atingindo seu ponto mínimo na sexta-feira (média de 1232 chegadas diárias).

Gráfico 21 - Resíduos



Fonte: elaborado pelo autor

Por fim, o gráfico de resíduos mostra que a série temporal possui um comportamento bem definido, visto que os valores residuais se encontram em intervalo absoluto de no máximo 200 chegadas, tendo sua média próxima de 0.

4. APLICAÇÃO DOS MODELOS DE PREVISÃO

Para a aplicação dos modelos de previsão, foi adotada uma abordagem de treinamento-teste, na qual 80% dos dados foram designados como conjunto de treinamento. Essa divisão permite que os modelos aprendam padrões a partir dos dados históricos (conjunto de treinamento) e, em seguida, sejam testados em 20% dos dados restantes para avaliar sua capacidade de generalização.

4.1. Média Móvel Simples

O primeiro método utilizado foi a média móvel simples, que é calculada tomando-se a média aritmética dos últimos n valores da série temporal. A fórmula para o cálculo da média móvel simples em um determinado ponto de tempo t é:

$$SMA = \frac{A_1 + A_2 + \dots + A_n}{n} \quad (17)$$

Onde:

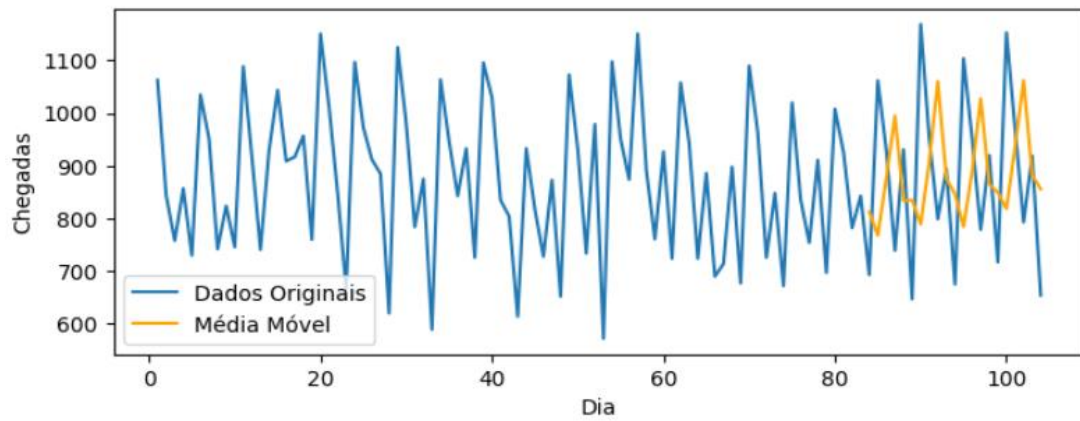
- SMA é o valor da média móvel simples no instante t ;
- A_n é a chegada real no período n ;
- n é o número de períodos.

Quanto maior o número de períodos, menor é a influência dos valores mais recentes para o cálculo da previsão. Para realizar os testes, foram escolhidos os períodos $n = 2$ e $n = 5$.

4.1.1. Período = 2

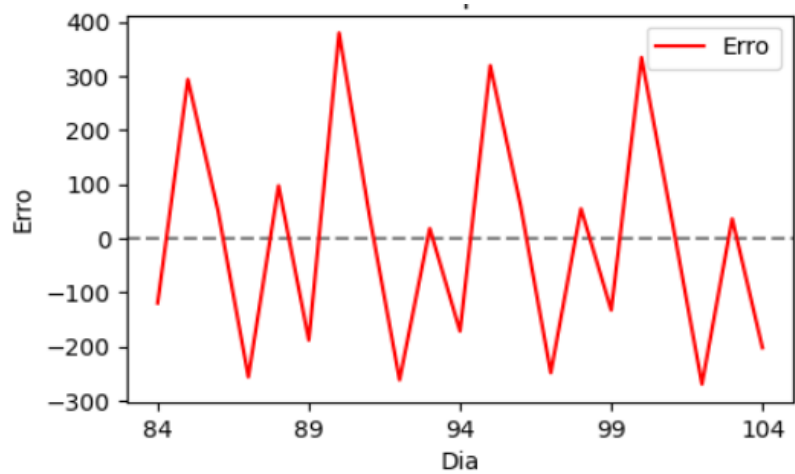
Primeiro, realizamos a previsão por média móvel com número n de períodos igual a 2, que considera as chegadas reais dos 2 dias anteriores.

Gráfico 22 - Média Móvel Simples (n = 2)



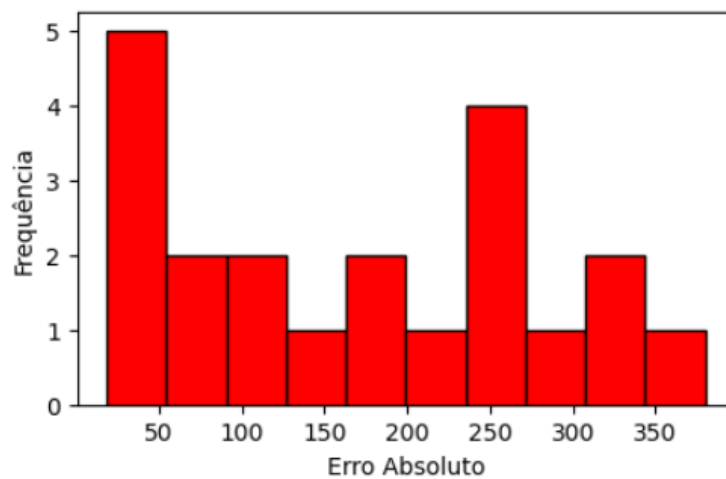
Fonte: elaborado pelo autor

Gráfico 23 - Erro - Média Móvel Simples (n = 2)



Fonte: elaborado pelo autor

Gráfico 24 - Histograma de erros absolutos - Média Móvel Simples (n = 2)



Fonte: elaborado pelo autor

Por ser um método pouco complexo, acaba não capturando a influência da forte sazonalidade presente nessa série temporal. Além disso, a previsão por média móvel tem um atraso inerente na resposta às mudanças nos padrões, resultando em elevados erros (como mostra o quadro abaixo):

Quadro 5 – Erros na previsão por Média Móvel (n = 2)

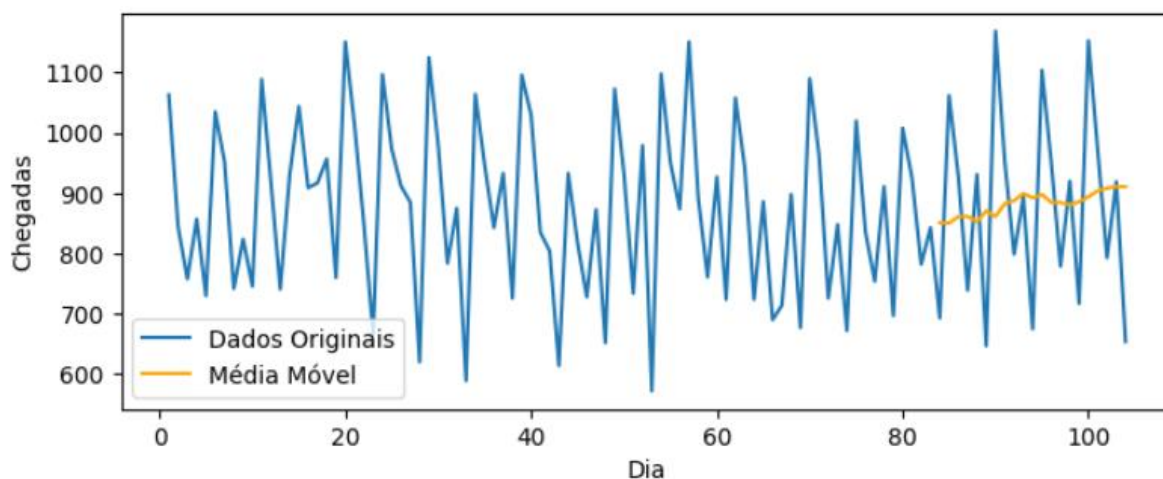
Erro	Valor
Erro Médio (ME)	-5.74
Erro Médio Absoluto (MAE)	170.36
Erro Percentual Médio Absoluto (MAPE)	19.78%
Raiz do Erro Quadrático Médio (RMSE)	204.05

Fonte: elaborado pelo autor

4.1.2. Período = 5

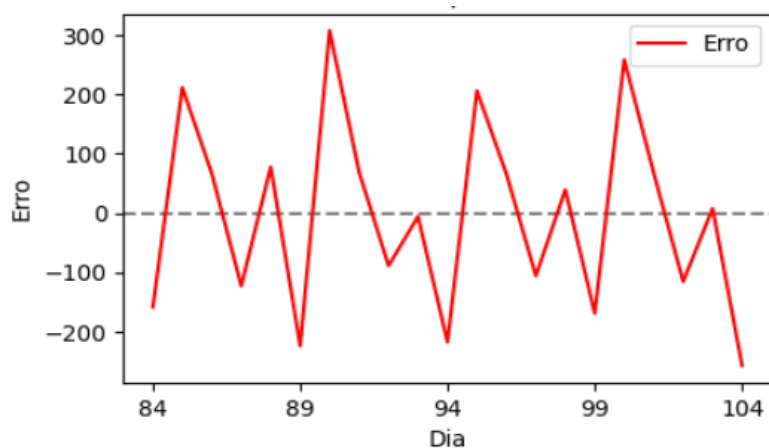
O próximo período testado é o de n = 5, ou seja, a previsão será realizada com base nas chegadas reais dos 5 dias anteriores.

Gráfico 25 - Média Móvel Simples (n = 5)



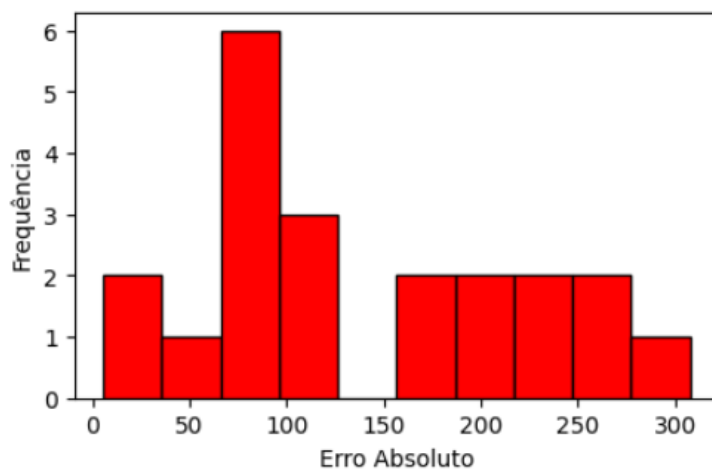
Fonte: elaborado pelo autor

Gráfico 26 - Erro - Média Móvel Simples (n = 5)



Fonte: elaborado pelo autor

Gráfico 27 – Histograma de erros absolutos - Média Móvel Simples (n = 5)



Fonte: elaborado pelo autor

A escolha de $n = 5$ produziu uma leve melhora na precisão. No entanto, ainda observamos erros elevados, pois acaba não capturando o padrão sazonal de chegadas.

Quadro 6 - Erros na previsão por Média Móvel (n = 5)

Erro	Valor
Erro Médio (ME)	-3.92
Erro Médio Absoluto (MAE)	135.33
Erro Percentual Médio Absoluto (MAPE)	16.10%
Raiz do Erro Quadrático Médio (RMSE)	160.18

Fonte: elaborado pelo autor

4.2. Regressão Linear

O segundo método de previsão utilizado é a regressão linear - técnica estatística utilizada para estabelecer uma relação linear entre uma variável dependente e uma ou mais variáveis independentes (ou de preditores). O modelo de regressão linear é representado pela seguinte equação:

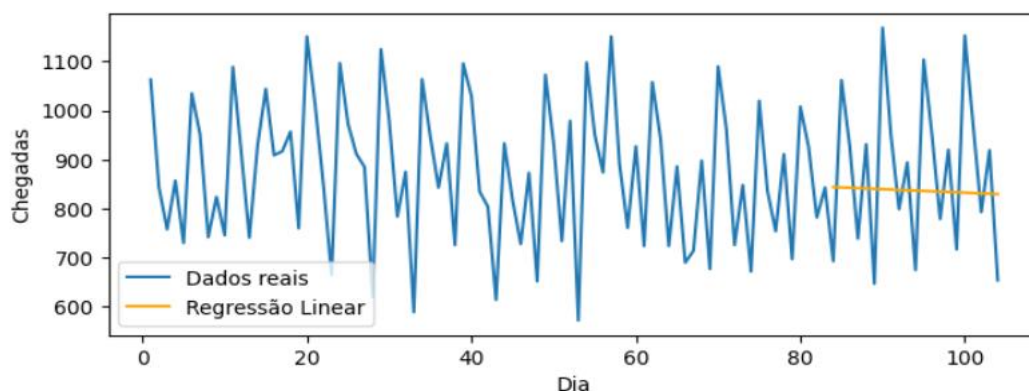
$$y = \beta_0 + \beta_1 * x_1 + \dots + \beta_k * x_k + \varepsilon \quad (18)$$

Onde:

- y é a variável dependente que desejamos prever;
- x_1, x_2, \dots, x_k são as variáveis independentes que podem afetar y ;
- $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ são os coeficientes de regressão que representam as inclinações da reta;
- ε representa o termo de erro, que é a diferença entre os valores observados e os valores previstos pelo modelo.

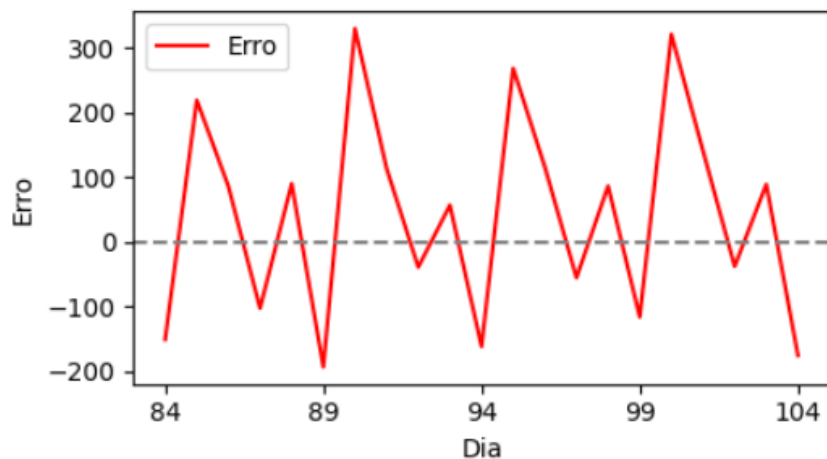
Para a aplicação da regressão linear, empregou-se a biblioteca “Scikit-Learn” no Python, fazendo uso da classe “LinearRegression”. Essa biblioteca proporciona uma implementação robusta e flexível de técnicas de regressão, permitindo a modelagem eficaz dos relacionamentos entre variáveis. Uma das vantagens do seu uso é a automação do processo de encontro de coeficientes, tornando mais eficiente a identificação dos pesos ótimos para cada variável independente.

Gráfico 28 - Regressão Linear



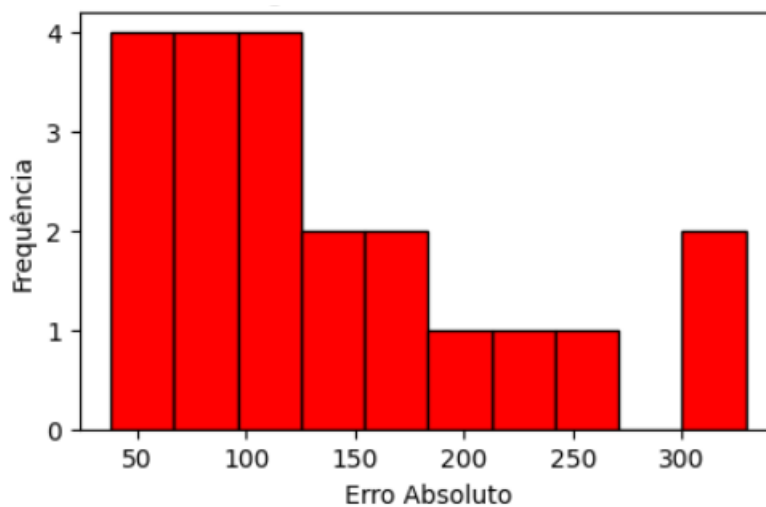
Fonte: elaborado pelo autor

Gráfico 29 - Erro - Regressão Linear



Fonte: elaborado pelo autor

Gráfico 30 – Histograma de erros absolutos - Regressão Linear



Fonte: elaborado pelo autor

A regressão linear pode não capturar adequadamente os padrões não lineares associados à sazonalidade. A modelagem linear é limitada em sua capacidade de representar flutuações cíclicas, levando a previsões imprecisas em períodos de variação sazonal, como é possível observar no quadro de erros abaixo:

Quadro 7 – Erros na previsão por Regressão Linear

Erro	Valor
Erro Médio (ME)	41.74
Erro Médio Absoluto (MAE)	140.26
Erro Percentual Médio Absoluto (MAPE)	15.86%
Raiz do Erro Quadrático Médio (RMSE)	162.96

Fonte: elaborado pelo autor

4.3. Suavização Exponencial Tripla (Holt-Winters)

A suavização exponencial tripla, também conhecida como Holt-Winters, é uma técnica avançada que leva em consideração a tendência e a sazonalidade, proporcionando uma modelagem mais precisa e abrangente. Essa abordagem é especialmente útil quando a série temporal apresenta padrões sazonais e tendências. Como as chegadas ao hospital possuem forte sazonalidade, é um modelo mais adequado do que os mencionados anteriormente, como média móvel e regressão linear. O modelo de suavização exponencial tripla é representado pelas seguintes equações:

$$B_t = \alpha * \frac{D_t}{I_{t-L}} + (1 - \alpha) * (B_{t-1} + T_{t-1}) \quad (19)$$

$$T_t = \beta * (B_t - B_{t-1}) + (1 - \beta) * T_{t-1}$$

$$I_t = \gamma * \frac{D_t}{B_t} + (1 - \gamma) * I_{t-L}$$

$$F_t(t + k) = (B_t + kT_t) * I_{t-L+k}$$

$$k = 1, 2, \dots, L$$

Equações de suavização exponencial tripla

Onde:

- D_t - demanda do período t;
- B_t - base ao final do instante t;
- T_t - tendência ao final do instante t;
- I_t - índice de sazonalidade do instante t;
- α - constante de suavização para Base;
- β - constante de suavização para Tendência;
- γ - constante de suavização para Sazonalidade;
- $F_t(u)$ - previsão ao final do período t para o período u ($u > t$);
- L - Tempo de uma estação completa.

Para realizar a suavização exponencial tripla, foi empregada a classe “ExponentialSmoothing” do módulo “statsmodels.api.tsa”. Uma das principais vantagens da utilização dessa classe é a capacidade de realizar um ajuste estatístico para seleção automática dos melhores parâmetros do modelo. Isso é crucial, pois elimina a necessidade de ajuste manual, permitindo que o próprio algoritmo identifique a combinação mais apropriada de tendência, sazonalidade e suavização.

Outro ponto a favor é a capacidade de explorar paralelização e otimização, o que contribui significativamente para a eficiência computacional, especialmente em grandes conjuntos de dados. A execução paralela acelera o processo de ajuste do modelo, tornando-o mais rápido e adequado para lidar com grandes quantidades de dados.

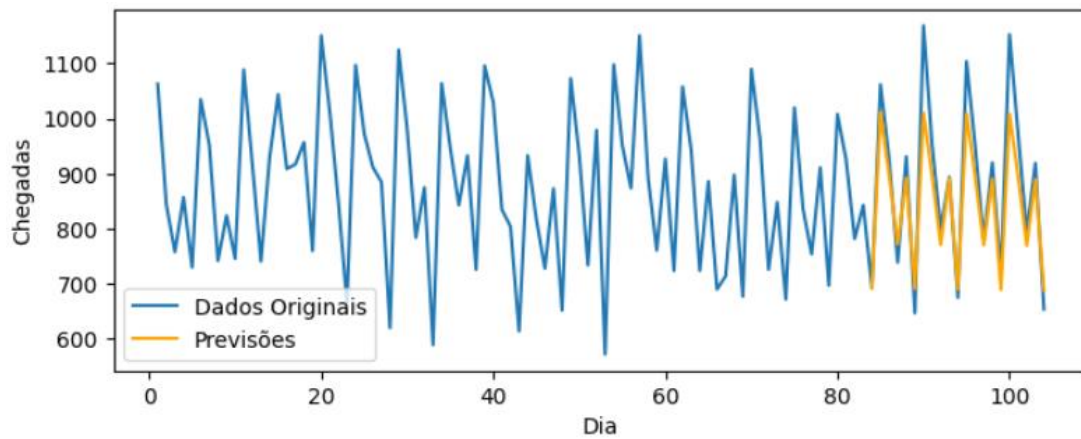
A tendência foi modelada como aditiva (“add”), enquanto a sazonalidade também foi considerada aditiva (“add”), com uma periodicidade sazonal de 5 dias por semana (“seasonal_periods” = 5). Essa abordagem visa capturar tanto as tendências de longo prazo quanto os padrões sazonais específicos, considerando a natureza semanal dos dados.

A escolha de modelar a tendência e a sazonalidade como aditivas (add) foi feita considerando o comportamento da série temporal. Em uma abordagem aditiva, as contribuições da tendência e da sazonalidade são adicionadas diretamente à série temporal. Isso implica que as variações nos dados são interpretadas como mudanças absolutas, o que se alinha ao comportamento observado no gráfico de tendência (gráfico 19) e no gráfico de sazonalidade (gráfico 20).

Na modelagem multiplicativa, as contribuições da tendência e da sazonalidade são multiplicadas à série temporal. Isso implica que as variações são interpretadas como mudanças proporcionais, o que seria apropriado se as flutuações relativas fossem mais relevantes em relação ao nível absoluto.

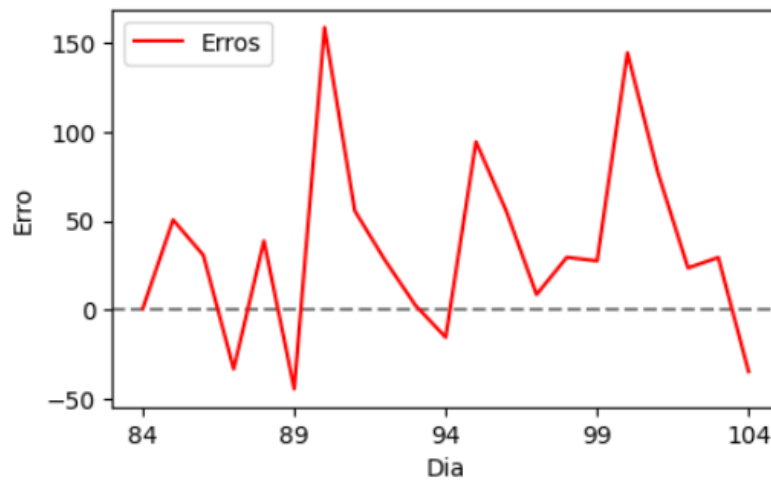
Portanto, a escolha por uma abordagem aditiva é apropriada visto que as variações absolutas na tendência e na sazonalidade são mais significativas para a análise do que as variações proporcionais.

Gráfico 31 - Suavização Exponencial Tripla



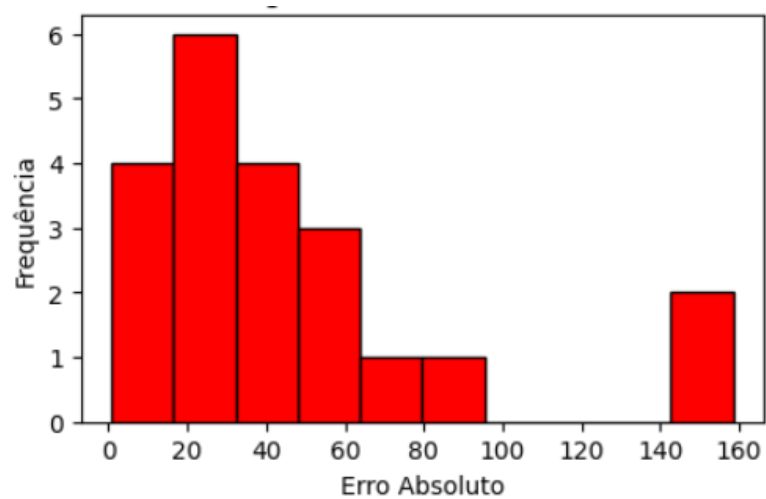
Fonte: elaborado pelo autor

Gráfico 32 - Erro - Suavização Exponencial Tripla



Fonte: elaborado pelo autor

Gráfico 33 - Histograma de erros absolutos - Suavização Exponencial Tripla



Fonte: elaborado pelo autor

O modelo retornou o parâmetro α igual a 0.005. Isso significa que o algoritmo está atribuindo uma importância relativamente baixa para as observações mais recentes ao fazer previsões. Em outras palavras, o modelo está dando menos peso à observação mais recente em comparação com as observações históricas ao calcular a previsão futura.

O parâmetro β também foi estimado como 0.005. Isso indica que a suavização da tendência no modelo é baixa, dando menos importância às tendências recentes ao fazer previsões futuras. O modelo está mais inclinado a confiar nas tendências de longo prazo em vez de reagir rapidamente a mudanças recentes.

Quanto ao parâmetro γ , seu valor é 0.64. Esse valor significa que o modelo atribui uma importância moderada à sazonalidade mais recente ao calcular as previsões sazonais. Ele está equilibrando entre a dependência das sazonalidades passadas e a capacidade de se ajustar a padrões sazonais mais recentes.

Ao adicionar a componente de sazonalidade no modelo, podemos observar uma expressiva queda nos erros da previsão (a variação do RMSE resultante da suavização exponencial tripla em relação ao RMSE resultante da regressão linear foi de -62%, aproximadamente):

Quadro 8 – Erros na previsão por Suavização Exponencial Tripla

Erro	Valor
Erro Médio (ME)	34.67
Erro Médio Absoluto (MAE)	46.79
Erro Percentual Médio Absoluto (MAPE)	4.93%
Raiz do Erro Quadrático Médio (RMSE)	61.85

Fonte: elaborado pelo autor

4.4. Modelo ARIMA

Como mencionado anteriormente, o Modelo ARIMA (AutoRegressive Integrated Moving Average) combina três componentes essenciais para modelagem de séries temporais: o componente autorregressivo (AR), a média móvel (MA) e a diferenciação (I). A equação do ARIMA (p,d,q) é expressa por:

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (20)$$

Equação do modelo ARIMA

Onde:

- p - ordem do componente autoregressivo (AR), que representa o número de lags incluídos no modelo;
- d - ordem de diferenciação, que representa o número de vezes que a série temporal foi diferenciada para torná-la estacionária;
- q - ordem do componente de média móvel (MA), que representa o número de lags dos erros incluídos no modelo;
- Y_t - é o valor da série temporal no tempo t ;
- ϕ_1, \dots, ϕ_p são os coeficientes autoregressivos;
- $\theta_1, \dots, \theta_q$ são os coeficientes da média móvel;
- ε_t - termo de erro no tempo t .

O componente AR mede como a observação atual da série temporal está relacionada às observações anteriores em termos de atrasos (ou lags). A ordem do componente AR é representada como "p" e indica quantos lags passados são considerados no modelo.

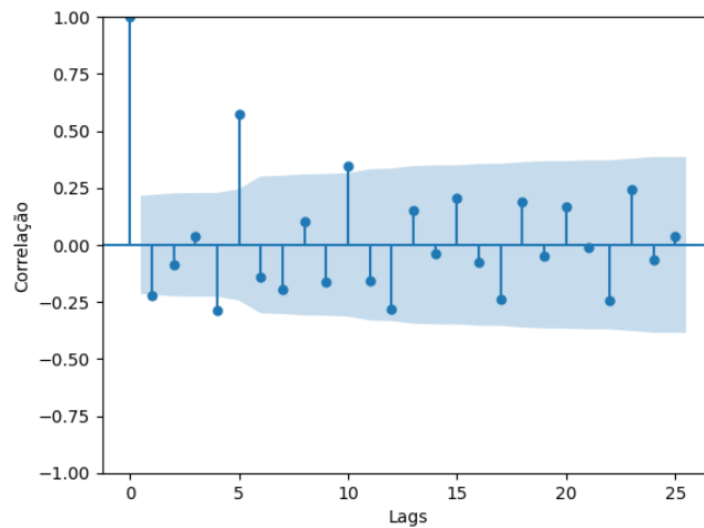
Já o componente MA modela como a observação atual depende dos erros (ou resíduos) passados em termos de atrasos (lags). A ordem do componente MA é representada como "q" e indica quantos erros passados são considerados no modelo.

Por fim, a diferenciação refere-se ao processo de tornar a série temporal estacionária, o que é necessário para aplicar o ARIMA. A ordem de diferenciação é representada como "d" e indica quantas vezes a série é diferenciada para atingir a estacionariedade. Isso geralmente envolve calcular a diferença entre uma observação e a observação anterior.

Para selecionar os parâmetros de maior precisão do modelo, é possível utilizar algumas métricas, como o AIC, o BIC ou o RMSE. O Critério de Informação Akaike (AIC) é uma métrica usada para comparar modelos estatísticos, incluindo modelos ARIMA, com o objetivo de escolher o modelo que melhor se ajusta aos dados e, ao mesmo tempo, evita a complexidade excessiva. Quanto menor o valor do AIC, melhor o modelo é considerado em termos de equilíbrio entre ajuste e complexidade.

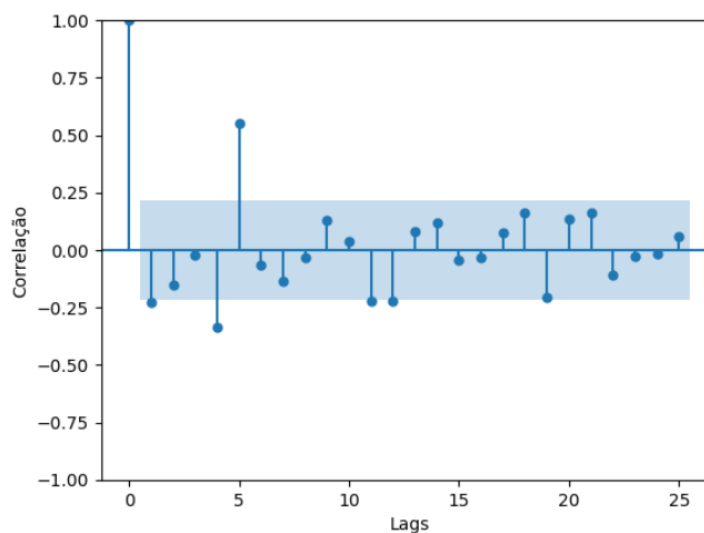
A fim de obter os melhores parâmetros do modelo, podemos utilizar os gráficos de autocorrelação (ACF) e de autocorrelação parcial (PACF):

Gráfico 34 - ACF



Fonte: elaborado pelo autor

Gráfico 35 - PACF



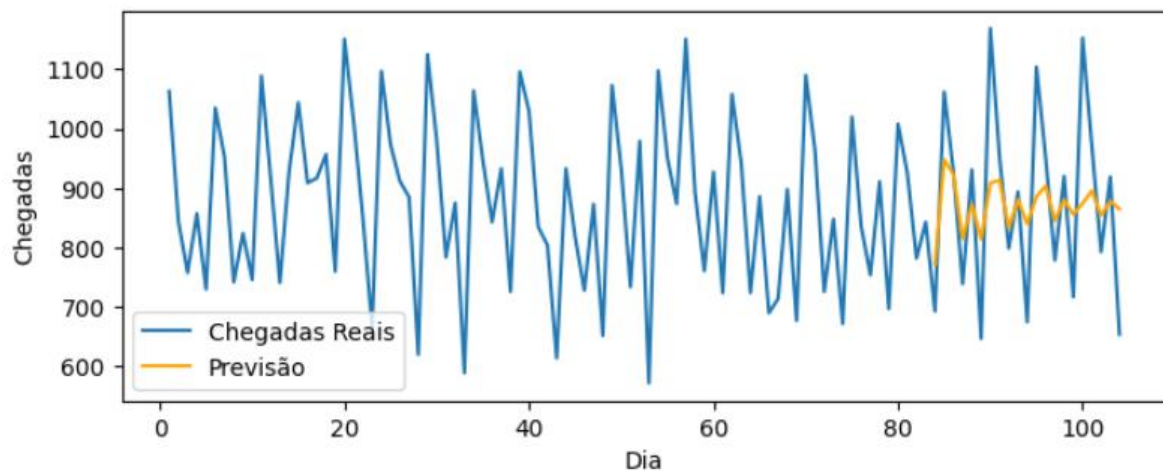
Fonte: elaborado pelo autor

Segundo o gráfico de autocorrelação, temos o lag 5 como significativo, visto que se encontra fora dos limites de confiança de 95%. Após esse lag, observamos um comportamento de senoide amortecida. Já em relação ao gráfico de autocorrelação parcial, também observamos o lag 5 como significativo, fora dos limites de confiança de 95%. Após esse lag, existe um corte súbito.

De acordo com Montgomery, Jennings e Kulahci (2015), esse padrão nos gráficos de ACF e PACF implica em um modelo AR (5). Visto que a série é estacionária, não é necessária diferenciação, ou seja, o parâmetro “d” é 0. Assim, a melhor combinação de parâmetros do modelo ARIMA (p,d,q) é (5,0,0), o qual produziu um AIC (métrica de desempenho) de 1032.

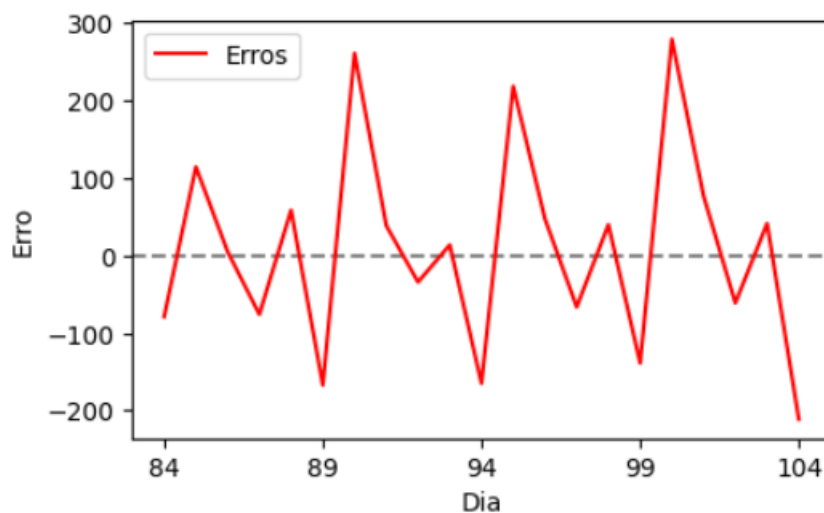
Para aplicar o modelo ARIMA, foi utilizada a classe “ARIMA” do módulo “statsmodels.tsa.arima.model” no Python, que gerou a seguinte previsão:

Gráfico 36 - Aplicação do modelo ARIMA (5,0,0)



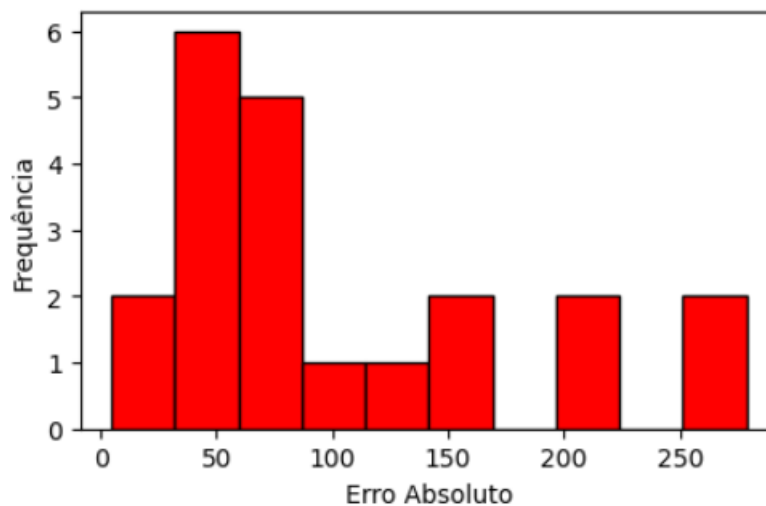
Fonte: elaborado pelo autor

Gráfico 37 - Erro - Aplicação do modelo ARIMA (5,0,0)



Fonte: elaborado pelo autor

Gráfico 38 – Histograma de erros absolutos - Aplicação do modelo ARIMA (5,0,0)



Fonte: elaborado pelo autor

Quadro 9 – Erros na previsão pelo Modelo ARIMA

Erro	Valor
Erro Médio (ME)	9.13
Erro Médio Absoluto (MAE)	104.24
Erro Percentual Médio Absoluto (MAPE)	12.14%
Raiz do Erro Quadrático Médio (RMSE)	131.39

Fonte: elaborado pelo autor

Ao analisar as previsões geradas pelo modelo ARIMA para a série temporal, é evidente que encontramos erros elevados, devido à presença de forte sazonalidade nos dados. O ARIMA, embora seja um modelo robusto para lidar com tendências e padrões de curto prazo, pode ter limitações quando confrontado com variações sazonais mais complexas.

Diante disso, uma alternativa é considerar modelos mais complexos, como o SARIMA, que são projetados para lidar explicitamente com componentes sazonais. O SARIMA incorpora elementos específicos para capturar a complexidade de padrões sazonais, oferecendo uma abordagem mais apropriada para séries temporais com sazonalidade forte.

4.5. Modelo SARIMA

O Modelo SARIMA (Seasonal AutoRegressive Integrated Moving Average) é uma extensão do ARIMA que leva em consideração a sazonalidade presente em muitas séries temporais. Além dos componentes AR, I e MA do ARIMA, o SARIMA também inclui três componentes sazonais adicionais: o componente Sazonal Autorregressivo (SAR), o componente Sazonal de Diferenciação (D), e o componente Sazonal de Média Móvel (SMA). A equação do SARIMA $(p,d,q)(P,D,Q,s)$ pode ser expressa por:

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t + \Phi_1 Y_{t-s} + \dots + \Phi_P Y_{t-Ps} + \Theta_1 \varepsilon_{t-s} + \dots + \Theta_Q \varepsilon_{t-Qs} \quad (21)$$

Equação do modelo SARIMA

Onde:

- p - ordem do componente autorregressivo (AR) para os lags não sazonais;
- d - ordem de diferenciação não sazonal;
- q - ordem do componente de média móvel (MA), para os lags dos erros não sazonais;
- P - ordem do componente autorregressivo (AR) para os lags sazonais;
- D - ordem de diferenciação sazonal;
- Q - ordem do componente de média móvel (MA), para os lags dos erros sazonais;
- s - período sazonal, que representa o número de observações em uma temporada completa;
- Y_t - é o valor da série temporal no tempo t ;
- ϕ_1, \dots, ϕ_p são os coeficientes autorregressivos não sazonais;
- $\theta_1, \dots, \theta_p$ são os coeficientes de média móvel não sazonal;
- Φ_1, \dots, Φ_p são os coeficientes autorregressivos sazonais;
- $\Theta_1, \dots, \Theta_p$ são os coeficientes da média móvel sazonal;
- ε_t - termo de erro no tempo t .

O componente Sazonal Autorregressivo (SAR) é similar ao componente AR, entretanto, captura a relação linear entre uma observação e suas observações passadas em lags sazonais. A ordem do componente SAR é representada por "P".

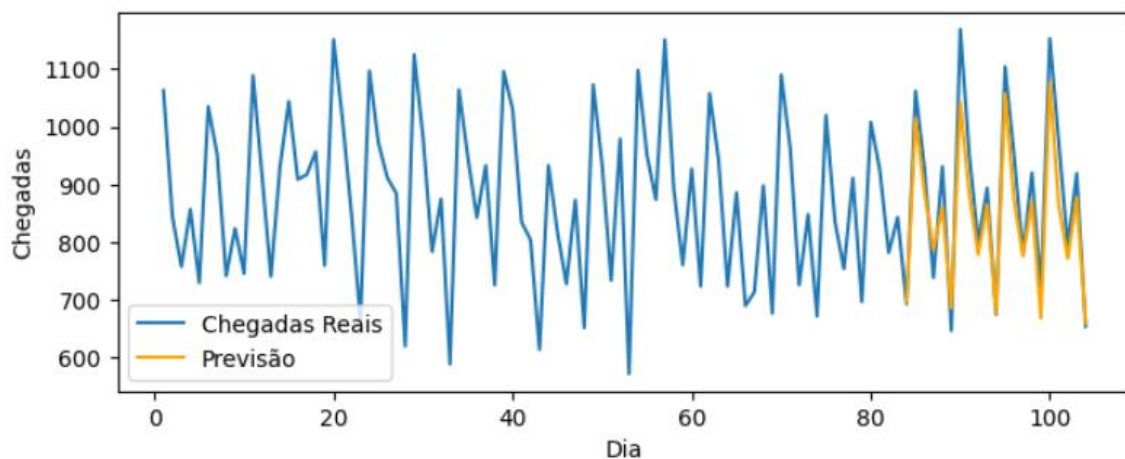
Já o componente Sazonal de Diferenciação (I) é análogo à diferenciação não sazonal, mas é aplicado em lags sazonais. A ordem do componente de diferenciação sazonal é indicada por "D".

Por fim, o componente Sazonal de Média Móvel (SMA) é similar ao componente MA, porém, captura a relação linear entre uma observação e os erros de previsão em lags sazonais. A ordem do componente SMA é denotada por "Q".

Para encontrar a melhor combinação de parâmetros P, D e Q para o modelo SARIMA (5,0,0)(P,D,Q,s), foi utilizado o método de "grid-search" no Python, com o critério AIC como métrica de precisão. O parâmetro s (número de observações em uma temporada completa) foi definido como 5, visto que o padrão de chegadas se repete a cada 5 dias (número de dias úteis na semana). O intervalo de valores de teste para os parâmetros P e Q foi definido como de 0 a 5, já o intervalo para o parâmetro D foi estipulado como de 0 a 2.

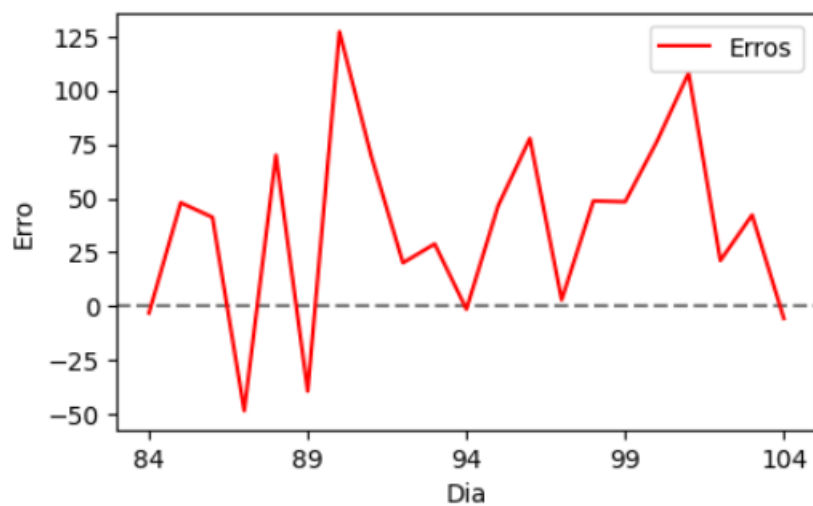
Como resultado, a melhor combinação gerada foi o modelo SARIMA (5,0,0)(0,2,1,5), que produziu um AIC de 943. Para aplicação do modelo, foi utilizada a classe "SARIMAX" do módulo "statsmodels.tsa.statespace.sarimax" no Python, que gerou a seguinte previsão:

Gráfico 39 - Aplicação do modelo SARIMA (5,0,0)(0,2,1,5)



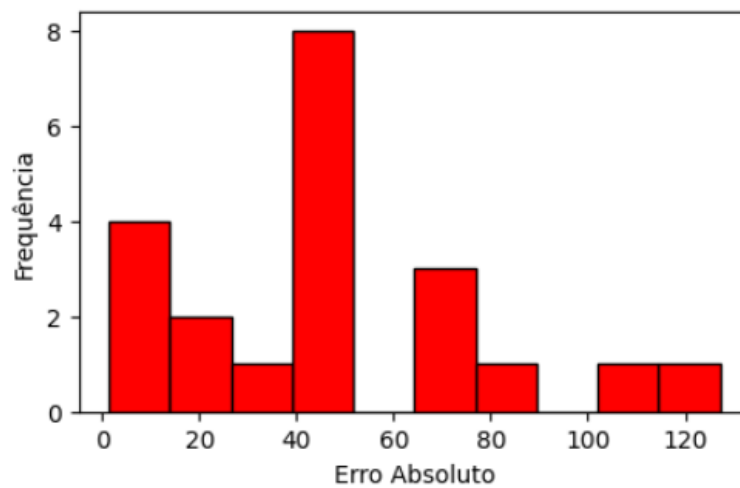
Fonte: elaborado pelo autor

Gráfico 40 - Erro - Aplicação do modelo SARIMA (5,0,0)(0,2,1,5)



Fonte: elaborado pelo autor

Gráfico 41 – Histograma de erros absolutos - Aplicação do modelo SARIMA(5,0,0)(0,2,1,5)



Fonte: elaborado pelo autor

Quadro 10 – Erros na previsão pelo Modelo SARIMA

Erro	Valor
Erro Médio (ME)	37.09
Erro Médio Absoluto (MAE)	46.42
Erro Percentual Médio Absoluto (MAPE)	4.97%
Raiz do Erro Quadrático Médio (RMSE)	56.8

Fonte: elaborado pelo autor

De acordo com o quadro acima, podemos observar uma expressiva queda nos erros, mostrando que o modelo SARIMA apresenta um desempenho superior, devido à sazonalidade forte na série temporal em questão.

4.6. Escolha do Melhor Modelo

Para escolher o modelo de maior precisão, utilizaremos os quatro tipos de erros apresentados – erro médio (ME), erro médio absoluto (MAE), erro percentual médio absoluto (MAPE), raiz do erro quadrático médio (RMSE) – associando pesos a cada um, de acordo com sua complexidade. Os pesos escolhidos foram:

- Erro Médio (ME): peso 1;
- Erro Médio Absoluto (MAE): peso 2;
- Erro Percentual Médio Absoluto (MAPE): peso 3;
- Raiz do Erro Quadrático Médio (RMSE): peso 4.

A fim de normalizar os cálculos, será utilizado o valor absoluto do erro médio, além de multiplicarmos por 100 o erro percentual médio absoluto.

Quadro 11 – Erros para cada modelo de previsão utilizado

Erro	Média Móvel	Regressão Linear	Holt-Winter	ARIMA	SARIMA
ME	-3.92	41.74	34.67	9.13	37.09
MAE	135.33	140.26	46.79	104.24	46.42
MAPE	16.10%	15.86%	4.93%	12.14%	4.97%
RMSE	160.18	162.96	61.85	131.39	56.8
Total	963.6	1021.68	390.44	779.59	372.04

Fonte: elaborado pelo autor

De acordo com o quadro acima, é possível observar que o modelo SARIMA é o mais adequado, pois possui o menor valor total com base no cálculo de erro utilizado – o que representa uma precisão mais elevada.

5. CONCLUSÃO

Ao concluir esta pesquisa, destaca-se a relevância do Instituto do Câncer do Estado de São Paulo (ICESP) como centro de referência no tratamento do câncer, enfrentando desafios operacionais decorrentes da variação na demanda de pacientes ao longo do tempo. Este estudo teve como objetivo realizar um comparativo entre modelos de previsão de demanda, buscando identificar a abordagem mais adequada para estimar com precisão as taxas de chegadas de pacientes e, assim, possibilitar um dimensionamento eficiente dos recursos hospitalares.

A revisão bibliográfica englobou os principais métodos de previsão, como suavização exponencial, modelos ARIMA e SARIMA, complementada por conceitos estatísticos relevantes. O mapeamento do processo de chegadas no hospital, realizado por meio de visitas de campo e análise de dados, proporcionou insights valiosos para a aplicação dos métodos de previsão.

Os resultados obtidos, ao aplicar os métodos de previsão em comparação com os dados reais, foram avaliados por métricas de erro, culminando na identificação do modelo mais adequado para otimizar o dimensionamento dos recursos disponíveis. Esta abordagem visa aprimorar a assertividade na previsão da demanda futura, contribuindo para a redução do tempo de espera dos pacientes.

A implementação do modelo escolhido tem o potencial de melhorar significativamente a eficiência operacional do ICESP, resultando em serviços mais ágeis, redução de filas e maior satisfação dos pacientes. Além disso, a disponibilização do código-fonte e o treinamento proposto para um integrante designado pelo instituto visam assegurar a continuidade e a autonomia na gestão do modelo, fortalecendo ainda mais seu impacto positivo no ambiente hospitalar.

Em síntese, este estudo não apenas fornece uma abordagem mais precisa para a previsão de demanda no ICESP, mas também oferece uma estratégia concreta para melhorar a qualidade dos serviços prestados, contribuindo para a eficiência global do atendimento no instituto.

6. PRÓXIMOS PASSOS

Após a avaliação de diversos modelos de previsão, incluindo média móvel, regressão linear, suavização exponencial tripla, ARIMA e SARIMA, identificamos que, por meio de métricas de precisão, o modelo SARIMA se destacou como a escolha mais eficaz para lidar com a forte sazonalidade presente nos dados de demanda hospitalar fornecidos pelo ICESP.

O próximo passo envolve a apresentação dos resultados e recomendações aos administradores do hospital. Pretendemos destacar a eficácia do modelo SARIMA na previsão de padrões sazonais e discutir os benefícios da sua implementação. Entendemos que a precisão do modelo é fundamental para otimizar a gestão de recursos e a eficiência operacional do hospital.

Contudo, reconhecemos uma limitação significativa associada à utilização de dados de 2019. Recomendamos uma avaliação da viabilidade de obter conjuntos de dados mais recentes, visando fortalecer ainda mais a robustez do modelo e garantir que as previsões reflitam as condições atuais e futuras do ICESP.

Além disso, uma vez implementado, o modelo SARIMA deve ser submetido a um treinamento contínuo, ajustando-se aos novos dados à medida que se tornam disponíveis. A integração eficiente com outros sistemas e ferramentas utilizadas no ICESP, como sistemas de gerenciamento de recursos e informações hospitalares, é essencial para garantir uma implementação bem-sucedida.

Estabelecer métricas de desempenho claras para avaliar o sucesso contínuo do modelo, incluindo métricas de precisão das previsões, é fundamental. Essa avaliação regular ajudará a identificar áreas de melhoria e garantirá que o modelo continue a agregar valor ao ICESP.

Adicionalmente, sugerimos explorar outras oportunidades para a aplicação do modelo SARIMA dentro do ICESP, considerando áreas adicionais onde a previsão precisa possa ser crucial para a tomada de decisões informadas.

Para assegurar a continuidade e o sucesso do modelo, propomos disponibilizar o código-fonte em Python utilizado na pesquisa para um integrante designado pelo hospital. Essa medida visa garantir a manutenção e o correto funcionamento do modelo no ambiente do ICESP. Adicionalmente, oferecemos a realização de um breve treinamento para o integrante designado, visando capacitá-lo na operação e

compreensão do código, permitindo que ele assuma a responsabilidade pela gestão contínua do modelo.

Por fim, o sucesso da implementação do modelo SARIMA no ICESP tem o potencial de aprimorar significativamente a capacidade de previsão e gestão de recursos, contribuindo para a otimização da eficiência operacional do hospital.

REFERÊNCIAS BIBLIOGRÁFICAS

icesp.org.br - Site do ICESP

Hyndman, R.J., & Athanasopoulos, G. (2021) "Forecasting: principles and practice", 3rd edition, OTexts: Melbourne, Australia

Makridakis, S., Wheelwright, S.C., & Hyndman, R.J. (1998). "Forecasting: methods and applications." 3rd Edition. Wiley.

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2015). "Time Series Analysis: Forecasting and Control." John Wiley & Sons.

Chatfield, C. (2004). "The Analysis of Time Series: An Introduction." Chapman and Hall/CRC.

Brockwell, P. J., & Davis, R. A. (2002). "Introduction to Time Series and Forecasting." Springer.

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). "Introduction to Linear Regression Analysis." Wiley.

Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). "Introduction to Time Series Analysis and Forecasting". John Wiley & Sons.

Shumway, R. H., & Stoffer, D. S. (2017). "Time Series Analysis and Its Applications: With R Examples." Springer.

Gardner, E. S., & McKenzie, E. (1985). "Forecasting trends in time series." Management Science.

Hyndman, R. J., & Khandakar, Y. (2008). "Automatic time series forecasting: The forecast package for R." Journal of Statistical Software.

<https://online.stat.psu.edu/stat510> - Penn State Eberney College of Science

Mehl, M. R., Conner, T. S. (2012) "Handbook of Research Methods for Studying Daily Life." Guilford Press.

Davis, R. A., & Hsu, S. L. (2008). "Introduction to Time Series Analysis." Springer.

Namhias, S. (2009). "Production and Operations Analysis." 6.ed., McGraw-Hill.