

Rosana Ribeiro da Costa

**Séries Temporais : análise das despesas da
Secretaria da Saúde do Estado do Rio Grande
do Sul antes e durante a pandemia do Covid-19**

Rio Grande, Rio Grande do Sul, Brasil

Fevereiro, 2022

Rosana Ribeiro da Costa

Séries Temporais : análise das despesas da Secretaria da Saúde do Estado do Rio Grande do Sul antes e durante a pandemia do Covid-19

Monografia apresentanda ao Instituto de Matemática, Estatística e Física da Universidade Federal do Rio Grande, como requisito para a obtenção do título de Bacharel em Matemática Aplicada.

Universidade Federal do Rio Grande - FURG

Instituto de Matemática, Estatística e Física - IMEF

Curso Bacharelado em Matemática Aplicada

Orientador: Prof^a Doutora Raquel da Fontoura Nicolette

Coorientador: Economista Kyze Quadros Mello

Rio Grande, Rio Grande do Sul, Brasil

Fevereiro, 2022

Rosana Ribeiro da Costa

Séries Temporais : análise das despesas da Secretaria da Saúde do Estado do Rio Grande do Sul antes e durante a pandemia do Covid-19

Monografia apresentanda ao Instituto de Matemática, Estatística e Física da Universidade Federal do Rio Grande, como requisito para a obtenção do título de Bacharel em Matemática Aplicada.

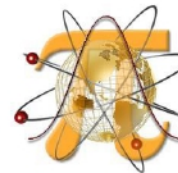
Trabalho aprovado.
Rio Grande, Rio Grande do Sul, Brasil
25 de Fevereiro de 2022.



Universidade Federal do Rio Grande – FURG

Instituto de Matemática, Estatística e Física
Curso de Bacharelado em Matemática Aplicada

Av. Itália km 8 Bairro Carreiros
Rio Grande-RS CEP: 96.203-900 Fone (53)3293.5411
e-mail: imef@furg.br Sítio: www.imef.furg.br



Ata de Defesa de Monografia

No vigésimo quinto dia do mês de fevereiro de 2022, às 14h, por webconferência, foi realizada a apresentação pública da defesa do Trabalho de Conclusão de Curso da acadêmica **Rosana Ribeiro da Costa**, sob orientação da Prof.^a Dr.^a Raquel da Fontoura Nicolette, deste instituto, e intitulada **Séries Temporais: análise das despesas da Secretaria da Saúde do Estado do Rio Grande do Sul antes e durante a pandemia do Covid-19**. Para participar da banca avaliadora junto à orientadora foram convidados o Prof. Dr. Eduardo André Tillmann – ICEAC/FURG e a Prof.^a Dr.^a Débora Spenassato - IMEF/FURG. Concluídos os trabalhos de apresentação e arguição, a candidata foi: (x) aprovada por unanimidade; () aprovada somente após satisfazer as exigências que constam na folha de modificações, no prazo fixado pela banca; () reprovada. Na forma regulamentar, foi lavrada a presente ata, que é abaixo assinada pelos membros da banca, na ordem acima relacionada.

Documento assinado digitalmente
gov.br RAQUEL DA FONTOURA NICOLETTE
Data: 25/03/2022 15:45:56-0300
Verifique em <https://verificador.itl.br>

Prof.^a Dr.^a Raquel da Fontoura Nicolette
Orientadora

Documento assinado digitalmente
gov.br EDUARDO ANDRE TILLMANN
Data: 25/02/2022 16:21:14-0300
Verifique em <https://verificador.itl.br>

Prof. Dr. Eduardo André Tillmann

Documento assinado digitalmente
gov.br DEBORA SPENASSATO
Data: 25/02/2022 15:54:33-0300
Verifique em <https://verificador.itl.br>

Prof.^a Dr.^a Débora Spenassato

Este trabalho é dedicado aos meus pais que não mediram esforços para que eu tivesse a oportunidade de prosseguir meus estudos. Sem o apoio deles não teria chegado até aqui, esse é apenas o começo de muitas vitórias que ainda estão por vir. .

Agradecimentos

Gostaria de agradecer primeiramente a Deus e a minha orientadora Raquel Fontoura Nicolette, pela sua disponibilidade e por aceitar o convite para me orientar fiquei muito honrada, também gostaria de agradecer-lhe por todo seu conhecimento e experiência compartilhados, mas principalmente por ser mais que uma orientadora, foi minha psicóloga no momento que pensei em desistir de tudo com toda sua paciência me mostrou que a vida sempre vamos ter obstáculos e por mais impossível que pareça ser de alguma forma chegamos lá. E também por ter me auxiliar na escolha do tema, sem ela não teria conseguido e por todas as aulas de inferência que fez eu me apaixonar e querer seguir nesta área tão linda que é a estatística.

A minha coorientadora Kyze Quadros Mello por toda paciência e estar disposta a contribuir com o meu trabalho, sei que meu perfeccionismo prejudicou um pouco, mas você sempre me traz para a realidade e diz que não estou em uma dissertação e sim em uma monografia. Ela tem razão, mas meu toc não deixa, todos os dias tento superar isso. Mas é a melhor e muito mais que um coorientadora, tornou-se uma amiga que vou levar para a vida.

Ao professor Eduardo e a professora Débora por ter aceitado o convite e compor a banca da minha monografia, com certeza vocês tem muito a contribuir com o meu trabalho. E ao professor Juliano, tenho um carinho muito especial por ele, esse cara foi meu professor de probabilidade na engenharia lembro até hoje de suas aulas, adorava explicar e eu anotar (haha) quando cheguei a ele estava sem norte, sabia que gostaria de trabalhar com séries temporais, mas não sabia a aplicação que utilizaria, então ele disse menina procure a professora Raquel, ela domina o assunto. Sou eternamente grata por ter feito a ponte para que o convite chegasse até ela.

A minha mãe Rosângela Ribeiro da Costa, por ter acreditado em mim desde pequena. Nunca fui um gênio da matemática, confesso que sempre tive dificuldade e na escola ia em exame quase sempre nessa disciplina, pois desde pequena estudava sozinha, meus pais trabalhavam e não tinham condições de me ajudar continuamente nos dever de casa. Lembro que sempre dizia mãe eu odeio matemática, nunca vou fazer um curso de exatas, e ela calmamente me dizia filha um dia serás professora de matemática e odiar é uma palavra muito forte, diz que tu ama matemática e verás que todas tuas dificuldades vai passar.

Durante o ensino médio comecei a gostar um pouquinho dela e ai resolvi fazer engenharia no 3º ano do curso vi que faltava algo que engenharia não era para mim, então resolvi mudar radicalmente e ir fazer um curso de matemática o resultado foi amar

o que faço e se tivesse a oportunidade faria novamente e no fim essa pessoa tinha razão e hoje sou muito grata a ela e se cheguei até aqui foi por ela. Mãe sou eternamente grata por sua vida e por sempre saber o que me dizer em todos os momentos.

Ao meu pai Sérgio Lavadouro da Costa que apesar de não ter jeito de falar as coisas sempre me apoiou de alguma forma, principalmente nos meus momentos de desespero. Dizem que somos muito parecidos, pois não sabemos nos expressar muito bem principalmente quando achamos que temos razão.

Ao meu irmão Rafael Ribeiro da Costa, por me ensinar a ter paciência (haha) principalmente quando foi necessário ajuda-lo nas tarefas da escola, confesso que muitas vezes eu queria sair correndo haha. Mas, me ensinasse muito principalmente a ser uma pessoa melhor e buscar ter mais conhecimento de alguns assuntos.

Ao meu amor Vagner da Silva Braga, que me acompanha nessa reta final de curso sabe o quanto é difícil e me dá todo o apoio necessário.

A minha amiga Ana Carolina, por todo apoio emocional principalmente quando eu pensava em desistir, lá estava ela me dando força e dizendo que chegaria lá.

E por último a todas as pessoas que passaram em minha vida e de alguma forma contribuíram nas minhas decisões que fizeram de alguma forma eu chegar até aqui.

A caminhada foi longa, mas hoje me sinto suficientemente madura para ver que todos os desvios que tive fez com que eu me fortalece-se e conseguisse chegar nos meus objetivos. Também gostaria de abrir um parênteses e agradecer a Universidade Federal do Rio Grande, por me proporcionar diversas experiências uma delas foi na área de pesquisa, quando fiz engenharia tive a oportunidade de realizar pesquisa no Hospital Universitário da FURG, lá fiz muitas amizades que levo comigo pelo resto de minha vida, conheci pessoas que me mostram que lugar de mulher é onde ela quiser e sim ela é capaz de ser mãe, pesquisadora e professora

*“Ensinar não é transferir conhecimento,
mas criar as possibilidades para a sua própria
produção ou a sua construção.”*
(Paulo Freire)

Resumo

As séries temporais constituem tópico importante para análise de dados econômicos e da saúde, principalmente em um período de pandemia global provocado pelo Sars-CoV-2 que impulsionou políticas públicas na área da saúde. Dessa forma, este trabalho teve como objetivo verificar se as despesas públicas da Secretaria de Saúde do Estado do Rio Grande do Sul (RS) no período compreendido entre o ano de 2018 e o primeiro semestre de 2021 podem ser descritos por uma série temporal e qual modelo que melhor se ajusta para descrever os dados e suas projeções. Os dados destas despesas públicas foram obtidos no Portal Transparência RS e analisadas no programa R utilizando a interface RStudio. Entre os modelos estudados, obteve-se o melhor ajuste para despesas de capital com ARIMA (2,0,1) e para despesas corrente com ARIMA (2,0,2). Em contrapartida, as estimativas com melhor ajuste para os modelos TAR resultaram em TAR (2,0,1) para despesas de capital e TAR (1,0,1) para as despesas correntes. Para ambos modelos o melhor ajuste foi determinado com base no critério de Akaike (AIC). Assim, ajustados os modelos de séries temporais, foi possível realizar a projeção para despesas correntes e de capital para os próximos 6 meses de investimento.

Palavras-chaves: Orçamento Público, Saúde Pública, Séries Temporais.

Resumo

Time series are an important topic for the analysis of economic and health data, especially in a period of global pandemic caused by Sars-CoV-2 that boosted public policies in the health field. Thus, this study aimed to verify whether the public expenditures of the Health Department of the State of Rio Grande do Sul (RS), in the period between 2018 and the first half of 2021, can be described by a time series and which model that best fits to describe the data and its projections. Data on these public expenditures were obtained from the Transparency Portal of RS and analyzed in the R program using the RStudio interface. Among the studied models, the best fit was obtained for capital expenditures with ARIMA (2.0.1) and for current expenditures with ARIMA (2.0.2). In contrast, the best-fit estimates for the TAR models resulted in TAR (2.0.1) for capital expenditures and TAR (1.0.1) for current expenditures. For both models, the best fit was determined based on the Akaike criterion (AIC). Thus, after adjusting the time series models, it was possible to project current and capital expenditures for the next 6 months of investment.

keywords: Public budget, Public health, Time Series.

Lista de ilustrações

Figura 1 – Funcionalidade dos órgãos responsáveis pelo planejamento orçamentário.	18
Figura 2 – Níveis de Gestão.	19
Figura 3 – Áreas da estatística	20
Figura 4 – Temperatura do ar em dado local no período de 24 horas	24
Figura 5 – Processo Estocástico de uma família de V.a	25
Figura 6 – Quadro com as especificação dos parâmetros utilizados	40
Figura 7 – Planilha dos parâmetros utilizados	41
Figura 8 – Representação gráfica da série temporal para despesa de capital	43
Figura 9 – Representação gráfica da série temporal para a despesa corrente	44
Figura 10 – Sazonalidade das despesas de capital	45
Figura 11 – Sazonalidade das despesas corrente	45
Figura 12 – Decomposição aditiva para a série de despesa de capital	46
Figura 13 – Decomposição multiplicativa para a série de despesa de capital	46
Figura 14 – Decomposição aditiva para a série de despesa de corrente	47
Figura 15 – Decomposição multiplicativa para a série de despesa corrente	48
Figura 16 – Função de autocorrelação para despesa de capital	49
Figura 17 – Função de autocorrelação para despesa corrente	49
Figura 18 – Remoção de Tendência para despesa de capital	52
Figura 19 – Remoção de Tendência para despesa corrente	52
Figura 20 – Previsão ARIMA para despesa de capital	55
Figura 21 – Previsão ARIMA para despesa corrente	56
Figura 22 – Representação gráfica do modelo TAR para despesa de capital	57
Figura 23 – Representação gráfica do modelo TAR para despesa corrente	58

Lista de tabelas

Tabela 1 – Investimento mensal em despesa corrente (log)	42
Tabela 2 – Estatística descritiva para despesa de corrente ao longo deste período (log)	42
Tabela 3 – Investimento mensal em despesa capital (log)	42
Tabela 4 – Estatística descritiva das despesas de capital (log)	42
Tabela 5 – Teste Box-Ljung	50
Tabela 6 – Estatística dos testes para despesa de capital	50
Tabela 7 – Estatística dos testes para despesa corrente	50
Tabela 8 – Teste de ajuste sazonal para despesa de capital	51
Tabela 9 – Teste de sazonalidade para despesa corrente	51
Tabela 10 – Ajuste ARIMA para despesa de capital	53
Tabela 11 – Ajuste ARIMA para despesa de corrente	53
Tabela 12 – Ajuste TAR para despesa de capital	54
Tabela 13 – Ajuste TAR para despesa corrente	54
Tabela 14 – Previsão pelo modelo ARIMA (2,0,1) para a despesa de capital	55
Tabela 15 – Previsão pelo modelo ARIMA (2,0,2) para a despesa de corrente	56
Tabela 16 – Ajuste da despesa de capital	57
Tabela 17 – Ajuste TAR para despesa corrente	58

Sumário

	Introdução	14
1	OBJETIVOS	16
1.1	Estrutura da Monografia	16
2	FUNDAMENTAÇÃO ECONÔMICA	17
2.1	Orçamento Público	17
2.2	Sistema de saúde no orçamento público do Rio Grande do Sul	17
2.3	Níveis de Gestão de acordo com a esfera de governo	18
2.4	Despesas Públicas	19
3	FUNDAMENTAÇÃO MATEMÁTICA	20
3.1	Conceitos básicos da estatística descritiva	20
3.1.1	Etapas da análise descritiva	21
3.1.2	Explorando os dados	21
3.1.2.1	Medidas de Centro	21
3.1.2.2	Medidas de Variação	22
3.1.2.3	Medidas de Posição Relativa	23
3.2	Série Temporal	23
3.3	Processo Estocástico	25
3.4	Linearidade de uma Série	26
3.5	Componentes das séries temporais	26
3.5.1	Tendência (T_t)	26
3.5.1.1	Modelo de Regressão Linear	26
3.5.1.2	Médias Móveis	29
3.5.1.3	Ajuste Exponencial	29
3.5.2	Sazonalidade	30
3.5.3	Ciclo	30
3.5.4	Resíduo	30
3.6	Decomposição da série	30
3.6.1	Decomposição Aditiva	30
3.6.2	Decomposição Multiplicativa	31
3.7	Estacionariedade da Série	31
3.7.1	Teste ADF	31
3.7.2	Teste PP	32
3.7.3	Teste KPSS	32

3.8	Autocorrelação	33
3.9	Função de Autocorrelação	33
3.10	Previsão	34
3.11	Modelos para séries temporais	34
3.11.1	Modelos de classe paramétrica	34
3.11.1.1	Box e Jenkins (1976)	34
3.11.1.2	Modelo- Autorregressivo AR(p)	35
3.11.1.3	ARMA (p,q)	35
3.11.1.4	ARIMA (p,d,q)	35
3.11.1.5	Modelo TAR	36
3.12	Crítérios para a escolha do melhor modelo	37
3.12.1	Crítério de Informação de Akaike (AIC)	37
4	APLICAÇÃO	39
4.1	Software R	39
4.2	Dados	39
4.3	Análise Descritiva	41
4.4	Análise de Séries Temporais	43
4.4.1	Decomposição	45
4.4.2	Estacionariedade	48
4.4.3	Ajuste do modelo	50
4.4.4	Previsão	55
5	CONSIDERAÇÕES FINAIS	59
	REFERÊNCIAS	60
	Apêndice A- Estimador de Máxima Verossimilhança	63
	Apêndice B - Código das Análises	64

Introdução

O sistema de saúde é um conjunto de relações políticas, econômicas e institucionais responsáveis pela condução dos processos referentes à saúde de uma dada população. Assim, cada país possui o seu próprio sistema de saúde com características específicas. No Brasil, o Sistema Único de Saúde (SUS) foi criado em 1988, assegurando a universalidade ao direito social a todos ao sistema de saúde pública (GIOVANELLA et al., 2012). Além disso, foram sancionadas a Lei Orgânica da Saúde nº 8.080/90 para regulamentar o SUS, bem como a Lei 8.142/90 para instituir a participação da comunidade na gestão do sistema e das transferências intergovernamentais de recursos financeiros. Apesar da promoção do acesso universal à saúde, como dever do estado, estar previsto em legislação constitucional e infraconstitucional, inúmeros desafios emergem para garantir um orçamento público que assegure esse direito social a todos (BRASIL, 2020).

Conforme dados disponíveis no Portal Transparência Estado do Rio Grande do Sul, os investimentos anuais em saúde giram na casa dos bilhões. Assim, somente no primeiro semestre do ano de 2018, o estado investiu cerca de 6,2 bilhões de reais na saúde. Contudo, no ano de 2020, com a advento da pandemia de covid-19, esses recursos giraram em torno de 6,8 bilhões de reais no primeiro semestre do ano. Apesar disso, no primeiro semestre de 2021 com avanço da vacinação e redução de ocupação de leitos hospitalares esses investimentos ficaram na casa de 5,1 bilhões de reais, representando uma redução em relação ao primeiro semestre de 2020.

Neste contexto, é importante realizar a comparação desses investimentos frente a modelos estatísticos, pois conforme definição de (SANTOS, 2019) a estatística é um ramo da ciência que a partir de informações oferece técnicas capazes de planejar, organizar, resumir, analisar e interpretar dados de um processo, podendo ainda trabalhar com elaboração de modelos para realizar previsões futuras.

Quando se trabalha com modelos matemáticos, a análise descritiva se torna importante para conhecer os dados estudados, pois a partir dela é possível determinar qual modelo melhor se adequa aos resultados (OLIVEIRA, 2011). Tendo essa resposta, justifica-se a utilização de uma série temporal para analisar o comportamento destes dados ao longo do tempo.

A partir da série temporal e das suas características é possível realizar um ajuste dos dados e como consequência é possível selecionar modelos com a finalidade de melhorar as projeções estatísticas. Dentre os modelos conhecidos, os mais estudados são: ARIMA, ARMA, e os modelos autorregressivo com limiar (TAR) por exemplo. Os melhores modelos ajustados serão utilizados para fazer uma projeção de quanto será gasto no prazo

de 6 meses.

Portanto, este estudo visa analisar a aplicação de recurso público do Estado do Rio Grande do Sul na área da saúde de 2018 ao primeiro semestre de 2021, com o objetivo principal de verificar qual modelo de série temporal melhor se ajusta para descrever os dados e suas projeções.

O presente estudo ganha notoriedade em um cenário pandêmico que matou milhares e contaminou milhões de pessoas somente no Rio Grande Sul. Assim, saber quanto foi aplicado e as projeções para o semestre seguinte destacam a relevância deste trabalho no estudo das políticas públicas de saúde no RS.

1 Objetivos

Objetivo Geral

Verificar se a aplicação de recurso público orçamentário do Estado do Rio Grande do Sul na área da saúde no período compreendido entre o ano de 2018 e o primeiro semestre de 2021 pode ser descrito por uma série temporal e qual modelo que melhor se ajusta para descrever os dados e suas projeções.

Objetivos Específicos

1. Descrever o comportamento e as propriedades das séries.
2. Ajustar os dados a um determinado modelo.
3. Analisar o investimento anual da Secretaria de saúde do RS.
4. Observar o aumento causado pela pandemia de COVID-19 nas despesas públicas com a saúde do RS.

1.1 Estrutura da Monografia

Para melhor compreensão do trabalho, a monografia está dividida como segue:

Capítulo 1 - Introdução: apresenta a motivação, justificativa e objetivos deste trabalho, assim com uma breve discussão sobre o Sistema Único de Saúde (SUS) e os investimentos realizado pelo Estado no setor.

Capítulo 2- Fundamentação econômica: aborda conceitos econômicos fundamentais para a compreensão deste estudo, subdividida nos seguintes tópicos: conceito de orçamento público, funcionamento do SUS no orçamento público, os níveis de gestão e por fim o conceito de despesa pública.

Capítulo 3- Fundamentação matemática: discute os conceitos de estatística básica incluindo estatística descritiva, análise exploratória de dados, conceito de séries temporais e os modelos estatísticos adotados para estas séries.

Capítulo 4- Aplicação: neste capítulo será discutido a metodologia de séries temporais aplicada neste estudo e os resultados dessa aplicação.

Capítulo 5- Considerações finais.

2 Fundamentação Econômica

O presente capítulo aborda alguns conceitos econômicos necessários para o entendimento do trabalho. Inicialmente, tem-se o conceito de orçamento público, seguido do funcionamento do sistema de saúde no orçamento público, os níveis de gestão de acordo com as esferas de governo e o conceito de despesa pública.

2.1 Orçamento Público

O orçamento público é uma das principais ferramentas que o gestor público dispõe para o cumprimento de suas funções. De acordo com (SANTOS, 2016) tais ferramentas são fundamentais para:

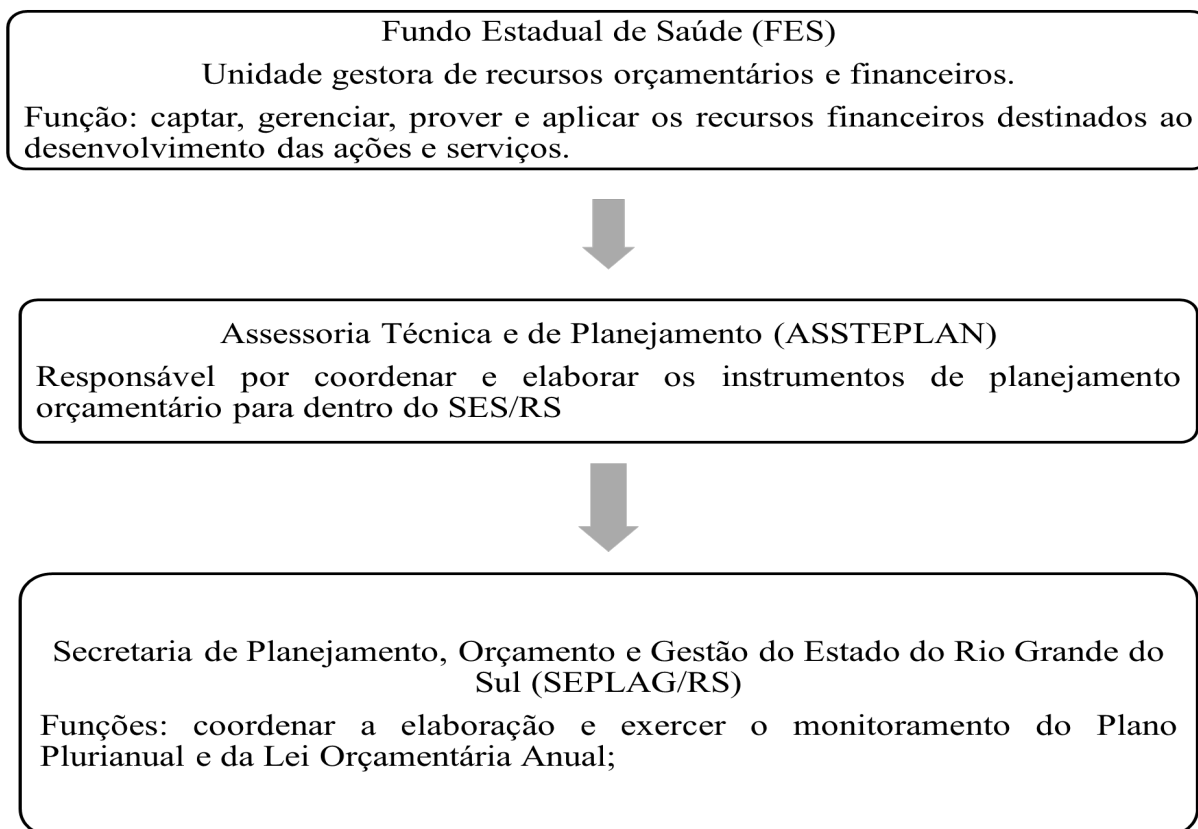
- Atender a demanda da comunidade por meio de bens e serviços;
- Suprir a demanda econômica voltada a atividades culturais, sociais e políticas da comunidade;
- Distribuir, por meio dos órgãos públicos competentes, o orçamento previsto e destinar o montante especificado e regulado por leis ou outras normas.

2.2 Sistema de saúde no orçamento público do Rio Grande do Sul

Através do planejamento orçamentário é realizada a previsão de aplicação dos recursos públicos, sob a competência dos órgãos do Fundo Estadual de Saúde (FES) juntamente com a Assessoria Técnica e de Planejamento (ASSTEPLAN) e com a Secretaria de Planejamento, Orçamento e Gestão do Estado do Rio Grande do Sul (SEPLAG/RS) (ACHUTTI, 2019). Cada órgão possui funcionalidades, conforme ilustrado na Figura 1.

O processo de elaboração do planejamento orçamentário é realizado com base na legislação federal, envolvendo diversas etapas, dentre as principais estão a aprovação da Lei do Plano Plurianual (PPA), a Lei de Diretrizes Orçamentárias (LDO) e da Lei Orçamentária Anual (LOA) (BRASIL, 2020). Cada etapa determina uma parte do planejamento: a LDO determina as metas e as prioridades do Plano Plurianual (PPA), que por sua vez, define as prioridades do Governo para o período de quatro anos, podendo ser revisado anualmente (BRASIL, 2020). A partir da LDO, é elaborada a LOA, que detalha todos os investimentos que serão realizados pelo governo: quanto será gasto, assim como a área de governo (saúde, educação, segurança pública) e a finalidade da aplicação (BRASIL, 2020).

Figura 1 – Funcionalidade dos órgãos responsáveis pelo planejamento orçamentário.



Fonte: Elaborada pelo autor.

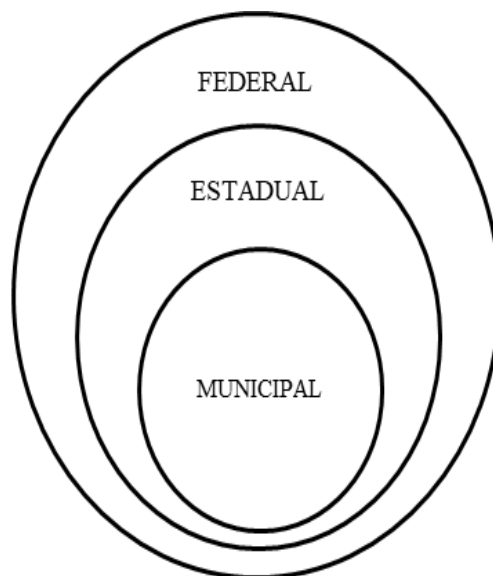
No estado do Rio Grande do Sul, a LOA compreende todas as receitas e despesas da administração direta e indireta. Enquanto a administração direta compreende a prestação de serviços públicos executados diretamente pelo próprio Estado e seus órgãos, a administração indireta consiste em todo o serviço prestado por pessoa jurídica criada pelo poder público para exercer determinada atividade (CUNHA, 2016). Neste contexto, o orçamento deve ser compatível com o Plano Plurianual (PPA) e deve estar em conformidade com a Lei de Diretrizes Orçamentárias (LDO) (TRANSPARÊNCIA, 2021).

2.3 Níveis de Gestão de acordo com a esfera de governo

A proteção e a defesa da saúde é competência concorrente da União, dos estados e dos municípios. Contudo, as atribuições políticas e financeiras variam conforme os níveis de gestão apresentados na Figura 2 (SANTOS, 2010). De acordo com (SANTOS, 2010), as esferas federal e estadual dispõem de um propósito comum relacionado à despesa pública no que se refere à aplicação e à distribuição dos recursos públicos.

A esfera federal é responsável por arrecadar os recursos e realizar a distribuição daquilo que é previsto para os estados e municípios. O estado também é responsável por destinar verbas públicas aos municípios. Contudo, o financiamento do SUS acontece

Figura 2 – Níveis de Gestão.



Fonte: Adaptada pelo autor.

de forma tripartite, ou seja, união, estados e municípios possuem vínculo de orçamento direcionado a saúde.

2.4 Despesas Públicas

A despesa pública é toda a aplicação do dinheiro resultante da arrecadação por meio de impostos ou outras fontes. Esses recursos custeiam o serviço público prestado à sociedade e os investimentos necessários para manutenção e/ou gestão do Estado (TRANSPARÊNCIA, 2021). Nos termos da Lei Federal nº 4.320/64, as despesas públicas podem ser classificadas em duas categorias econômicas, quais sejam:

(i) despesa corrente, aquela que está relacionada às despesas de custeio, aplicadas na manutenção e no funcionamento do serviço público e

(ii) despesa de capital, aquela que está associada ao investimento de capital para contribuir na geração de novos bens ou serviços que passarão a integrar o serviço público (BRASIL, 2020).

No ano de 2018, as despesas de capital do Rio Grande do Sul foram de aproximadamente 106 milhões de reais enquanto que as correntes foram em torno de 12 bilhões de reais. Já no ano de 2020, com o advento da pandemia de COVID-19, causada pelo vírus SARS-Cov-2, estas despesas atingiram 135 milhões em relação a despesa de capital, e 16 bilhões para a despesa corrente. Esse aumento decorreu do avanço da pandemia, da escassez de material e da falta de leitos de UTI'S (TRANSPARÊNCIA, 2021).

3 Fundamentação Matemática

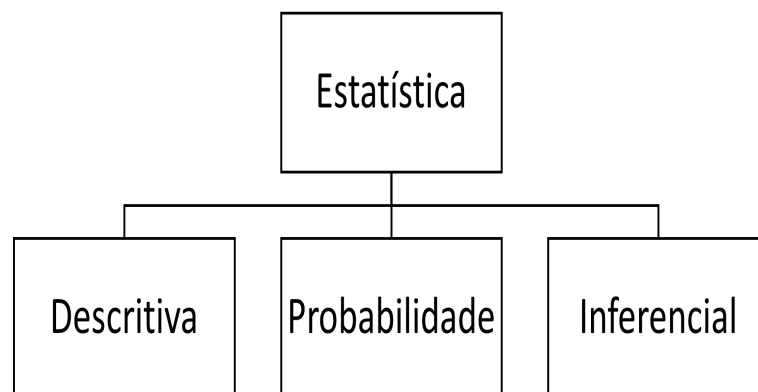
A fundamentação matemática e estatística é de grande importância na análise de dados, por conseguinte são fundamentais para a compreensão dos dados das despesas públicas do Estado do Rio Grande do Sul.

Este capítulo aborda conceitos fundamentais de estatística descritiva dos dados, os quais possibilitam uma análise exploratória dos dados, bem como definição matemática e conceitos específicos de análise de séries temporais.

3.1 Conceitos básicos da estatística descritiva

A estatística é um ramo da matemática que visa oferecer um conjunto de técnicas capaz de planejar, organizar, resumir, analisar e interpretar dados de um processo a partir de determinadas informações (SANTOS, 2019). Assim sendo, a estatística pode ser subdividida em três áreas nos termos da Figura 3.

Figura 3 – Áreas da estatística



Fonte:Elaborada pelo autor.

Cada uma das áreas possui sua especificidade neste estudo, por exemplo, dá-se ênfase inicial à análise descritiva e segue-se para a inferência com a construção de modelos autorregressivos para avaliação dos dados.

3.1.1 Etapas da análise descritiva

A análise descritiva é o ponto de partida do estudo, tendo em vista que a partir dela é possível definir o melhor método aplicável ao caso concreto (OLIVEIRA, 2011).

As etapas que compõem a análise descritiva são as seguintes:

- Definição do problema;
- Coleta de dados;
- Preparação e exploração dos dados;
- Análise exploratória de dados;
- Apresentação dos resultados.

3.1.2 Explorando os dados

Explorar os dados é fundamental para descrever as características da amostra em estudo (TRIOLA, 2017). Para isso foram utilizadas medidas de centro, medidas de variação e de posição relativas.

3.1.2.1 Medidas de Centro

A medida de centro é um valor central do conjunto de dados (TRIOLA, 2017). Existem diversas formas de se determinar o centro de um conjunto de dados.

- Média: A média ou simplesmente média aritmética, é uma medida de centro determinada através do somatório dos valores da amostra dividido pelo tamanho da amostra (TRIOLA, 2017). A equação 3.1 representa o cálculo da média.

$$\bar{x} = \frac{\sum_{i=1}^N x_i}{N} \quad (3.1)$$

onde:

\bar{x} - representa a média amostral;

$\sum_{i=1}^N x_i$ - somatório dos valores da amostra;

N - tamanho da amostra;

- Mediana: é uma medida de centro obtida através do valor do meio quando os dados originais estão arranjados em ordem crescente ou decrescente de magnitude (TRIOLA, 2017).

- Moda: é o valor que ocorre com maior frequência no conjunto de dados (TRIOLA, 2017).
- Ponto Médio: é a medida de centro do meio caminho entre os valores máximo e mínimo do conjunto (TRIOLA, 2017). O ponto médio pode ser calculado a partir da equação 3.2.

$$PM = \frac{VM + Vm}{2} \quad (3.2)$$

onde:

PM - Ponto médio;

VM - Valor de máximo;

Vm - Valor de mínimo;

3.1.2.2 Medidas de Variação

As medidas de variação são formas de quantificar e medir a variabilidade do conjunto de dados a ser estudado (TRIOLA, 2017).

- Amplitude: é a diferença entre o maior e o menor valor da amostra (TRIOLA, 2017). A equação 3.3. mostra o cálculo para se obter a amplitude em um conjunto de dados.

$$Amplitude = VM - Vm \quad (3.3)$$

onde:

VM - maior valor do conjunto de dados;

Vm - menor valor do conjunto de dados;

- Variância : é uma medida de variação do desvio quadrático médio da média (TRIOLA, 2017). Utilizando a equação 3.4, obtem-se a variância.

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{N - 1} \quad (3.4)$$

Onde:

s^2 - variância amostral;

x_i - valor dos elementos do conjunto de dados, $i = 1 \dots N$;

- Desvio Padrão : O desvio padrão é uma medida de variação que mede o grau de dispersão do conjunto de dados em relação a média (TRIOLA, 2017). O desvio padrão é calculado a partir da equação 3.5.

$$s = \sqrt{s^2} \quad (3.5)$$

onde:

s - desvio padrão;

s^2 - variância;

3.1.2.3 Medidas de Posição Relativa

As medidas de posição relativa indicam a posição do conjunto de dados em relação aos demais valores do mesmo conjunto de dados (TRIOLA, 2017).

- Escore z

O escore z é o número de desvios-padrão que um determinado valor assume abaixo ou acima da média (TRIOLA, 2017). O escore z é calculado a partir da equação 3.6.

$$z = \frac{x - \bar{x}}{s} \quad (3.6)$$

onde:

x - elementos amostrais;

\bar{x} - média;

s - desvio padrão;

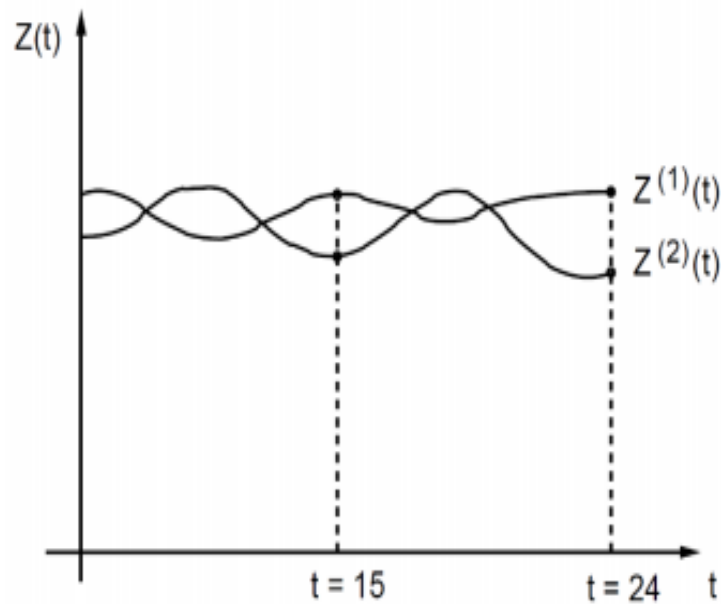
- Quartil

Os Quartis são medidas que dividem o conjunto de dados em quatro grupos com cerca de 25% dos valores em cada grupo (TRIOLA, 2017). Os quartis são denotados por Q_1, Q_2, Q_3, Q_4 .

3.2 Série Temporal

Uma série temporal é definida como um conjunto de observações ordenadas em um intervalo regular de tempo ao longo de um determinado período. Esse conjunto pode ser obtido através de observações periódicas, dentre as quais pode-se elencar: valores mensais de temperatura do ar (Figura 4); índices diários da Bolsa de Valores; número médio anual de manchas solares (MORETTIN; TOLOI, 2018).

Figura 4 – Temperatura do ar em dado local no período de 24 horas



Fonte: MORETTIN e TOLOI, 2018.

A Figura 4, exemplifica uma aplicação do conceito de séries temporais, pois é possível observar um conjunto de observações Z_t ao longo do tempo.

Na análise de uma série temporal, primeiramente deve-se modelar o problema a ser estudado, partindo da modelagem do problema é necessário descrever seu comportamento para fazer uma estimativa e por último, deve-se avaliar os fatores que influenciam no comportamento da série, buscando definir relações entre duas ou mais séries.

Segundo (SOEIRO et al., 2012), o que se conhece por série temporal é parte de uma trajetória dentre muitas outras que podem ser observadas. Denomina-se trajetória a curva obtida a partir do gráfico da série.

Assim, trajetória é um processo físico que está sendo observado, podendo ser expresso através de um vetor $\mathbf{Z}(\mathbf{t})$ representado na Equação 3.7.

$$\mathbf{Z}(t) = [Z_1(t), Z_2(t), Z_3(t)] \quad (3.7)$$

É importante ressaltar que para cada \mathbf{t} constante tem-se um valor de variável aleatória $\mathbf{Z}(\mathbf{t})$, que significa ter certa distribuição de probabilidade (MORETTIN; TOLOI, 2018). O conjunto de todas as possíveis trajetórias da série é denominado processo estocástico.

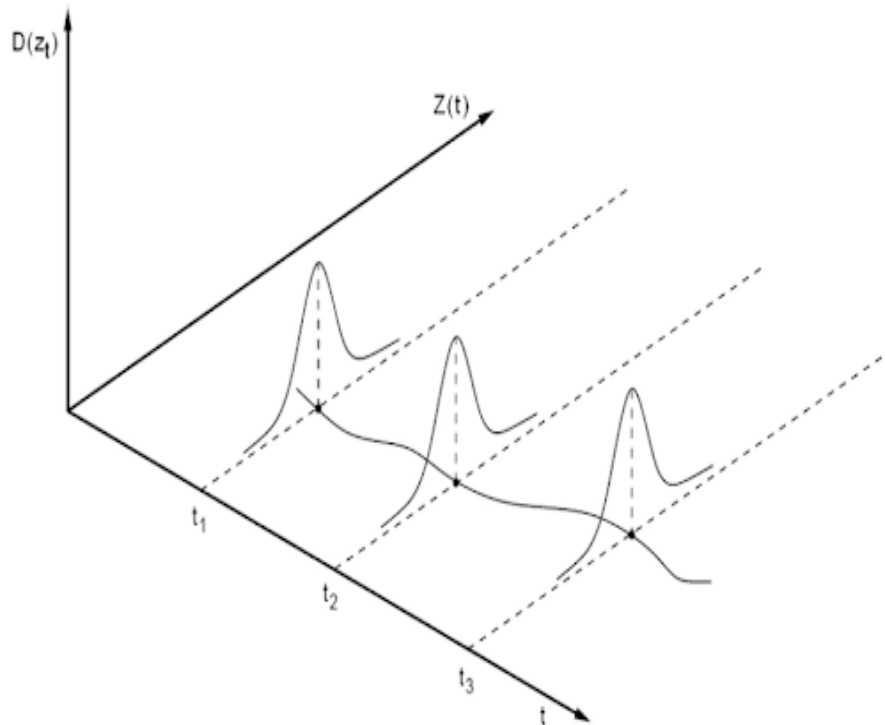
3.3 Processo Estocástico

Para (MORETTIN; TOLOI, 2018) o processo estocástico consiste em uma família, $Z = \{Z(t), t \in T\}$ tal que, para cada $t \in T$, $Z(t)$ é uma variável aleatória (v.a). Desta forma, entende-se que o processo estocástico é uma família de variáveis aleatórias definidas no mesmo espaço de probabilidade (Ω, P, A) .

Dessa forma, sendo Ω um espaço amostral, P a medida de probabilidade de um dado evento ocorrer e A o evento, pode-se então definir o espaço de probabilidade como uma tripla formada por (Ω, P, A) .

Ressalta-se que o conjunto T geralmente é tomado no conjunto dos números inteiros $\mathbb{Z} = \{0, \pm 1, \dots\}$ ou conjunto dos reais \mathbb{R} . Além disso, para cada $t \in T$, $Z(t)$ é uma v.a, definida sobre espaço amostral (Ω) . Então temos que a função $Z(t, w)$ com $t \in T$ e $w \in \Omega$ é de dois argumentos. A Figura 5 ilustra o processo estocástico (MORETTIN; TOLOI, 2018).

Figura 5 – Processo Estocástico de uma família de V.a



Fonte: MORETTIN e TOLOI, 2018.

Por fim, o processo estocástico pode ser classificado em parâmetro discreto ou contínuo.

3.4 Linearidade de uma Série

De acordo com BOSQUE (2014), a classificação de uma série linear pode ser escrita como:

$$Z_t = \mu \sum_{i=1}^n \phi a_{t-i} \quad (3.8)$$

Onde:

μ - média do processo Z_t ;

i - são números reais;

a_{t-i} - sequência de v.a;

Assim, qualquer série temporal que não possa ser escrita da forma da equação (3.8) é classificada como não-linear. O estudo de uma série temporal é repleto de modelos não lineares, dentre eles destacam-se o modelo ARCH, modelo ARMA e modelo autorregressivo com limiar (TAR).

3.5 Componentes das séries temporais

Uma série temporal pode ser composta por: tendência, sazonalidade, ciclo e resíduo (variável aleatória).

3.5.1 Tendência (T_t)

A tendência é uma componente da série temporal responsável por descrever o comportamento da variável em estudo ao longo do tempo e que se mantém ao remover as componentes sazonais e os residuais (REIS, 2020). Assim, ela pode ser calculada por meio de diversos modelos. Neste trabalho são utilizados os modelos de regressão linear, médias móveis e ajuste exponencial.

3.5.1.1 Modelo de Regressão Linear

O modelo de regressão linear é uma ferramenta utilizada entre economistas e matemáticos para descrever como a variável dependente Z_t está relacionada com as variáveis independentes t e com o termo do erro ϵ , com o objetivo de encontrar um modelo matemático que justifique essa relação (CHEIN, 2019). Assim, o modelo de regressão linear para a série de tendência é descrito pela Equação 3.11.

$$Z_t = \beta_0 + \beta_1 t + \epsilon \quad (3.9)$$

onde:

Z_t - variável dependente;

β_0 - coeficiente que intercepta o eixo y;

β_1 - coeficiente que define a inclinação da reta;

t - variável independente;

ϵ - erro padrão;

Os coeficientes β_0 e β_1 podem ser estimados pelo método de mínimos quadrados. Para realizar o método, pode-se utilizar os seguintes passos:

1 - Calcular a diferença entre Z_t e o valor estimado \widehat{Z}_t , ou seja, $(Z_t - \widehat{Z}_t)$

2 - Calcular o quadrado da diferença $(Z_t - \widehat{Z}_t)^2$

3 - Calcular o somatório das diferenças quadráticas $S = \sum (Z_t - \widehat{Z}_t)^2$

4º Para cada um dos coeficientes, deve-se aplicar as deveridas parciais da soma dos quadrados da diferença e igualar a zero para encontrar, onde a diferença se torna mínima.

Retornando a equação (3.9), vamos utilizar o método do mínimos quadrados para encontrar os coeficientes β_0 e β_1 .

Suponha, que os passos 1 e 2 tenham sido calculados, então partindo de 3, temos:

$$S = \sum_{i=1}^n (Z_t - \widehat{Z}_t)^2 \quad (3.10)$$

Substituindo $(\widehat{Z}_t)^2$ em (3.11), a equação pode ser reescrita como:

$$S = \sum_{i=1}^n (Z_t - \beta_0 - \beta_1 t_i)^2 \quad (3.11)$$

Os candidatos a ponto de mínimo da equação 3.12 são aqueles que as derivadas parciais são nulas de S em relação a cada um de seus parâmetros, isto é:

$$\frac{\partial S}{\partial \beta_0} = -2 \sum_{i=1}^n (Z_t - \beta_0 - \beta_1 t_i) = 0 \quad (3.12)$$

$$\frac{\partial S}{\partial \beta_1} = -2 \sum_{i=1}^n t_i (Z_t - \beta_0 - \beta_1 t_i) = 0 \quad (3.13)$$

Vamos dividir em duas partes. Primeiro realizando a manipulação matemática da equação 3.13

Ao multiplicarmos cada uma das variáveis pelo somatório, obtemos:

$$-2 \left(\sum_{i=1}^n Z_t - \sum_{i=1}^n \beta_0 - \sum_{i=1}^n \beta_1 t_i \right) = 0 \quad (3.14)$$

Posteriormente, dividimos a equação por 2 para simplificar a constante que aparece.

$$\frac{-2(\sum_{i=1}^n Z_t - \sum_{i=1}^n \beta_0 - \sum_{i=1}^n \beta_1 t_i)}{2} = 0 \quad (3.15)$$

Logo, temos:

$$-\sum_{i=1}^n Z_t + n\beta_0 + \sum_{i=1}^n t_i \beta_1 = 0 \quad (3.16)$$

Segundo, para realizar a manipulação da equação 3.14, seguimos as seguintes etapas:

Multiplicamos cada uma das variáveis pelo somatório

$$-2\left(\sum_{i=1}^n t_i Z_t - \sum_{i=1}^n \beta_0 t_i - \sum_{i=1}^n t_i^2 \beta_1\right) = 0 \quad (3.17)$$

Dividimos por 2 para simplificar com a constante que aparece.

$$\frac{-2(\sum_{i=1}^n t_i Z_t - \sum_{i=1}^n t_i \beta_0 - \sum_{i=1}^n t_i^2 \beta_1)}{2} = 0 \quad (3.18)$$

Logo:

$$-\sum_{i=1}^n t_i Z_t + \sum_{i=1}^n t_i \beta_0 + \sum_{i=1}^n t_i^2 \beta_1 = 0 \quad (3.19)$$

Assim, obtemos um sistema de equações normais, onde os coeficientes que procuramos calcular é β_0 e β_1 .

$$\begin{cases} n\beta_0 + \beta_1 \sum_{i=1}^n t_i = \sum_{i=1}^n Z_t \\ \beta_0 \sum_{i=1}^n t_i + \beta_1 \sum_{i=1}^n t_i^2 = \sum_{i=1}^n t_i Z_t \end{cases}$$

Escrevendo o sistema na forma matricial, tem-se:

$$\begin{vmatrix} n & \sum_{i=1}^n t_i \\ \sum_{i=1}^n t_i & \sum_{i=1}^n t_i^2 \end{vmatrix} \begin{vmatrix} \beta_0 \\ \beta_1 \end{vmatrix} = \begin{vmatrix} \sum_{i=1}^n Z_t \\ \sum_{i=1}^n t_i Z_t \end{vmatrix}$$

Resolvendo a matriz, obtem-se os resultados:

$$\beta_1 = \frac{n \sum_{i=1}^n t_i Z_t - \sum_{i=1}^n t_i Z_t}{n \sum_{i=1}^n t_i^2 - (\sum_{i=1}^n t_i)^2} \quad (3.20)$$

$$\beta_0 = \frac{\sum_{i=1}^n Z_t - \sum_{i=1}^n t_i \beta_1}{n} \quad (3.21)$$

Substituindo na equação original os valores de β_0 e β_1 , encontramos o valor para a tendência da série. Assim, observa-se que o modelo apresentado é um ajuste linear simples. Ressalta-se que existem alguns casos com séries de despesas que utilizam o ajuste linear multivariado, pois a variável Z_t é dependente de duas ou mais variáveis (PEDROSA, 2021). Desta forma, o modelo linear para este tipo de série é descrito pela equação a seguir:

$$Z_t = \beta_0 + \beta_1 t_1 + \beta_2 t_2 \dots + \beta_n t_n \quad (3.22)$$

3.5.1.2 Médias Móveis

Outra forma de obter a tendência é através das médias móveis. Segundo OLIVEIRA; CARVALHO (2013), as médias móveis consistem em calcular a média aritmética das n observações de um período anterior ao observado. Matematicamente pode ser calculado utilizando a equação:

$$M_t = \frac{\sum Z_{t-i}}{n} \quad (3.23)$$

onde:

M_t - Média móvel;

Z_t - série temporal que está sendo observada;

i - instantes anteriores ao período observado;

n - período;

3.5.1.3 Ajuste Exponencial

O ajuste exponencial é uma outra forma de calcular a tendência da série. Conforme REIS(2020), esse método apresenta algumas vantagens em relação ao das médias móveis, tais quais:

- Realizar previsões em curto período de tempo.
- Considerar todos os valores observados ao longo do período, o que não ocorre com as médias móveis que utilizar apenas valores próximo aos observados.

Para encontrar o ajuste exponencial e calcular a tendência, basta utilizar a equação:

$$E_i = W + Y_i + (1 - W)E_{i-1} \quad (3.24)$$

Onde:

i -um período qualquer de tempo;

Y_i -valor inicial da série no período i ;

E_i - série exponencialmente ajustada no período i ;

E_{i-1} - valor da série exponencialmente ajustada no período anterior;

W - Coeficiente de ajuste ($0 < W < 1$);

3.5.2 Sazonalidade

A componente sazonalidade em uma série temporal se refere a quando a série apresenta um comportamento periódico, oscilando de forma regular ao longo do tempo (RELVAS, 1998).

3.5.3 Ciclo

O ciclo de uma série temporal se refere aos movimentos recorrentes de elevação e queda em torno da tendência (LAMOUNIER, 2007).

3.5.4 Resíduo

A componente Ruído é um processo randômico no qual a instabilidade e a aleatoriedade não são explicados por variações cíclicas ou pela tendência, mas sim em função da presença de fatores aleatórios (KEILA, 2015).

3.6 Decomposição da série

A decomposição de uma série temporal permite dividir uma única série temporal em várias outras séries temporais diferentes para assim identificar quais componentes estão atuando naquele conjunto específico de dados. Geralmente, as séries temporais são divididas em 3 componentes: tendência, sazonalidade e o resíduo (REIS, 2020). Cada um destes componentes foram descritos no item 3.5 do presente capítulo. O procedimento de decomposição varia de acordo com o modelo de estudo e pode se subdividir em duas: decomposição aditiva ou decomposição multiplicativa.

Segundo HYNDMAN; ATHANASOPOULOS (2018) decomposição aditiva é apropriada se a variação sazonal ou a variação em torno do ciclo da tendência não varia de acordo com o nível da série temporal. Caso essas variações ocorram proporcionalmente de acordo com o nível da série, a decomposição multiplicativa é a mais indicada. Dessa forma, esse tipo de decomposição também é muito utilizado em séries econômicas.

3.6.1 Decomposição Aditiva

É a soma dos componentes tendência, sazonalidade e resíduo e o resultado da soma apresenta a mesma unidade das variáveis (REIS, 2020). A equação (3.26) representa esta decomposição.

$$Y = T_t + S_t + R_t \quad (3.25)$$

Onde:

Y - Resultado da decomposição;

T_t - Tendência;

S_t - Sazonalidade;

R_t - Resíduo (v.a.).

3.6.2 Decomposição Multiplicativa

É realizada através da multiplicação dos componentes tendência, sazonalidade e resíduo e o resultado da multiplicação apresenta a mesma unidade das variáveis (REIS, 2020). A equação (3.27) representa esta decomposição.

$$Y = T_t \times S_t \times R_t \quad (3.26)$$

Onde:

Y - Resultado da decomposição;

T_t - Tendência;

S_t - Sazonalidade;

R_t - Resíduo (v.a.).

3.7 Estacionariedade da Série

Uma série temporal é dita estacionária quando o conjunto de dados possui média e variância constantes ao longo do tempo, enquanto que a covariância deve indicar que a autocorrelação entre dois valores de uma variável y qualquer, tomados a partir de dois períodos de tempo diferentes, deve depender apenas do intervalo de tempo entre esses dois valores (FRACARO, 2018). Assim, para determinar a estacionariedade de uma série, basta aplicar alguns testes como: o teste ADF, o teste PP e o teste KPSS.

3.7.1 Teste ADF

Conforme (FRACARO, 2018), o teste ADF, também conhecido como Augmented Dickey-Fuller é um teste estatístico de hipótese nula, que tem por objetivo analisar as variáveis e verificar se estas são estacionárias, utilizando o teste de raiz unitária por meio da regressão a seguir:

$$\Delta Y_t = \alpha + \beta_t + \gamma Y_t - 1 + \sum_{i=1}^{p-1} \delta_i \Delta Y_t + a_t \quad (3.27)$$

Onde:

ΔY_t - é o operador das diferenças;

α - constante;

β_t - componente de tendência do modelo;

$\gamma Y_t - 1$ - testa a estacionaridade da série;

$\sum_{i=1}^{p-1} \delta_i \Delta Y_t$ - Defasagem;

a_t - Erro do Resíduo;

A partir da equação 3.29, são testadas as seguintes hipóteses:

$H_0 : \delta_i = 0.05$ e $H_1 : \delta_i < 0.05$. O teste é dado pela equação:

$$\tau = \frac{\hat{\delta}}{EP(\hat{\delta})} \quad (3.28)$$

Onde:

$EP(\hat{\delta})$ - erro padrão;

Por meio da análise de τ pode-se definir se a série dependente Y_t é estacionária ou não.

Se $|\tau|_{calculado} > |\tau|_{crítico}$, rejeita-se H_0 , então a série é estacionária.

Se $|\tau|_{calculado} < |\tau|_{crítico}$, não rejeita H_0 , então a série não é estacionária.

3.7.2 Teste PP

O teste de Phillips - Perron (PP) consiste em um procedimento não paramétrico presente em diversas classes de séries temporais em que existe raiz unitária (COSTA, 2016). Assim, o teste é descrito através da seguinte regressão:

$$\Delta Y_t = +\beta_1 + \beta_2 t + \delta Y_{t-1} + \sum_{i=1}^{p-1} \alpha_i \Delta Y_{t-i} + a_t \quad (3.29)$$

Da mesma forma que o teste anterior será necessário verificar τ para definir se a série dependente Y_t é estacionária ou não.

Se $|\tau|_{calculado} > |\tau|_{crítico}$, rejeita-se H_0 , então a série é estacionária.

Se $|\tau|_{calculado} < |\tau|_{crítico}$, não rejeita H_0 , então a série não é estacionária.

3.7.3 Teste KPSS

O teste KPSS - Kwiatkowski, Philips Schmidt e Shin - é um teste que avalia a estacionariedade da série considerando as componentes tendência, passeio aleatório e erro

(SILVEIRA; MATTOS; KONRATH, 2016). Esse teste é descrito pela equação (3.31).

$$Y_t = \xi t - r_t + a_t \quad (3.30)$$

Onde: ξ - constante r_t - passeio aleatório; a_t - Erro do Resíduo; Ao contrário dos testes ACF e PP, o teste KPSS, considera uma série estacionária:

Se $|\tau|_{calculado} < |\tau|_{crítico}$, rejeita-se H_0 , então a série é estacionária.

Se $|\tau|_{calculado} > |\tau|_{crítico}$, não rejeita H_0 , então a série não é estacionária.

3.8 Autocorrelação

A autocorrelação é uma ferramenta matemática que busca encontrar um padrão de repetição com a presença de um sinal periódico obscurecido por resíduo entre os valores da série em um determinado período de tempo e os valores da mesma série em outro período (LIMA et al., 2015).

3.9 Função de Autocorrelação

As funções de autocorrelações são conhecida pela sigla ACF e ajudam a caracterizar a série temporal Z_t ao longo do tempo. Além disso, são também utilizadas para fazer análise sobre regressões e encontrar raízes unitárias (LIMA et al., 2015). Essas funções podem ser expressas por:

$$\hat{p}_k = \frac{\sum_{t=k+1}^T [(x_t - \hat{\mu}) \times (x_{t-k} - \hat{\mu})]}{\sum_{t=k+1}^T (x_t - \hat{\mu})^2} \quad 0 < t \leq T - 1 \leq \hat{p}_k \leq 1 \quad (3.31)$$

Onde:

o numerador $\sum_{t=k+1}^T [(x_t - \hat{\mu}) \times (x_{t-k} - \hat{\mu})]$ corresponde a covariância de x_t em relação a média estimada;

O denominador $\sum_{t=k+1}^T (x_t - \hat{\mu})^2$ corresponde a variância de x_t em relação a média estimada;

Assim, os gráficos das funções de autocorrelação permitem identificar o número de termos autorregressivos.

Já as funções de autocorrelação parciais, também conhecidas como PACF, fornecem a correlação variável no período t e uma de suas defasagens, retirado os efeitos das demais defasagens (LIMA et al., 2015). A equação para essas funções é descrita por:

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + a_t \quad (3.32)$$

Onde:

o parâmetro β_2 fornece os coeficientes de correlação entre Y_t e Y_{t-2} ;

3.10 Previsão

As séries temporais são responsáveis por descrever o comportamento dos dados ao longo do tempo, além disso ela consiste em realizar previsões para valores futuros (BOSQUE, 2014). Essas previsões são baseadas em investigar a análise intrínseca de dados históricos para variáveis de interesse, com base nisso, são aplicados modelos liminar capazes de realizar essas previsões.

3.11 Modelos para séries temporais

Os modelos para séries temporais podem ser divididos em duas classes, de acordo com o número de parâmetros envolvidos: paramétricos ou não-paramétricos. É importante destacar que o objeto de estudo deste trabalho está restrito às classes paramétricas. Uma das metodologias mais utilizadas para a análise dos modelos paramétricos é a abordagem desenvolvida por (BOX; JENKINS, 1976). Dentre estes modelos, os mais frequentemente utilizados são os modelos de erro, o auto-regressivo e de média móvel (ARMA) e o modelo auto-regressivo integrado e de média móvel (ARIMA). Na classe não-paramétrica os modelos mais utilizados são a função de auto-covariância e sua transformada de Fourier (MORETTIN; TOLOI, 2018).

3.11.1 Modelos de classe paramétrica

3.11.1.1 Box e Jenkins (1976)

A metodologia (BOX; JENKINS, 1976) consiste em realizar um ajuste para modelos autorregressivos integrados de médias móveis ARIMA(p,d,q) a um conjunto de dados. A construção da estrutura do modelo é baseada nos próprios dados. As etapas para esta construção são:

I- Especificação : Uma classe geral dos modelos é considerada para análise, ou seja, os modelos ARIMA.

II- Identificação: A identificação do modelo ocorre com base nas análises de autocorrelação parcial e demais critérios.

III- Estimação: Nesta fase os parâmetros do modelo identificados são estimados.

IV- Verificação: Esta última etapa consiste em verificar o ajuste por meio de uma análise de resíduo e dizer se o modelo escolhido é adequado para o estudo.

3.11.1.2 Modelo- Autorregressivo AR(p)

O modelo autorregressivo, baseia-se na suposição de estacionaridade da série, utilizando a variável em estudo para fazer projeção futura a partir da variável anterior (CAVALARO, 2021). Assim, o modelo autorregressivo de ordem p é descrito por:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + a_t \quad (3.33)$$

Onde:

$X_{t-1} + X_{t-2} + \dots + X_{t-p}$ - são os p termos anterior a série X_t ;

$\phi_1 + \phi_2 + \dots + \phi_p$ - são os parâmetros reais da série;

a_t - Erro do Resíduo;

E a função de autocorrelação para o modelo autorregressivo é dada por:

$$\rho_t = \phi_1 \rho_{t-1} + \phi_2 \rho_{t-2} + \dots + \phi_p \rho_{t-p} \quad (3.34)$$

Onde:

$\rho_{t-1} + \rho_{t-2} + \dots + \rho_{t-p}$ são as correlações da variável X_t com as mesmas variáveis ao longo do período;

$\phi_1 + \phi_2 + \dots + \phi_p$ - são os parâmetros reais da série;

3.11.1.3 ARMA (p,q)

É um modelo autorregressivo e de médias móveis conhecido pela sigla (ARMA (p,q)). Segundo GLOT; BENGIO (2010), este modelo representa a combinação dos modelos AR e MA, onde a série Z_t é descrita por seus valores anteriores e pelo resíduo corrente passado. A equação 3.36, representa o modelo ARMA(p,q):

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \epsilon_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (3.35)$$

Onde:

ϕ_i = são os coeficientes da estrutura autorregressiva com $i = 1, 2, \dots, p$;

θ_i = são os coeficientes das estrutura de médias móveis, com $i = 1, 2, \dots, q$;

a_t = resíduo;

3.11.1.4 ARIMA (p,d,q)

Segundo GOMES (2016), os modelos de ARIMA (Auto-Regressive Integrated Moving Average Model), são modelos de classes mais gerais que realizam previsão para uma

série temporal estacionária. É importante destacar que nas séries não estacionárias, esse modelo pode ser aplicado desde que a série não estacionária seja transformada em estacionária por meio do método de diferenciação.

As variáveis que compõem esse modelo são:

- p - ordem da parte autorregressiva;
- d- é o número de diferenças não sazonais necessárias para a estacionaridade;
- q- ordem da parte de médias móveis;

Logo, esses modelos são uma generalização dos modelos ARMA. A equação que descreve o modelo ARIMA para uma série temporal estacionária é uma equação linear, descrita por:

$$\widehat{Y}_t = \mu + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (3.36)$$

Onde:

μ - média;

ϕ_i = são os coeficientes da estrutura autorregressiva com $i = 1, 2, \dots, p$;

θ_i = são os coeficientes das estrutura de médias móveis, com $i = 1, 2, \dots, q$;

Y_{t-1} - diferenciação;

e_{t-1} - resíduo;

3.11.1.5 Modelo TAR

O modelo Threshold Autoregressive Model (TAR), também chamado de modelo autorregressivo com limiar, inicialmente foi proposto por Tong no ano de 1978 e extensivamente estudado por Tong e Lim (1980). O objetivo desse modelo é aproximar uma série não-linear para uma série linear por partes, sendo assim esse modelo pode ser escrito como uma composição de modelos autorregressivos, como mostra a Equação 3.38 (BOSQUE, 2014).

$$y_t = \begin{cases} \phi_0 + \sum_{i=1}^{k_1} \phi_i y_{t-i} + a_t(1) \\ \phi_0 + \sum_{i=1}^{k_2} \phi_i y_{t-i} + a_t(2) \end{cases}, \text{ se } y_t \in \mathbb{R} \quad (3.37)$$

Onde:

ϕ_0 - operador autorregressivo;

y_{t-i} - defasagem;

a_t - Erro do Resíduo;

3.12 Critérios para a escolha do melhor modelo

Para avaliar qual modelo melhor explica o comportamento dos dados é imprescindível utilizar critérios para determinar o mínimo de parâmetros possíveis de serem estimados. Assim, dentre os critérios mais utilizados para selecionar o modelo estão: o Critério de Informação de Akaike (AIC), Critério Bayesiano de Schwarz (BIC), o critério de Hannan e Quinn (HQC), o critério Final Predictor Error (FPE) e o critério Criterion Autoregressive Transfer Function (CAT) (MORETTIN; TOLOI, 2018). Neste trabalho, o interesse é utilizar o Critério de Informação Akaike.

3.12.1 Critério de Informação de Akaike (AIC)

Segundo MORETTIN; TOLOI (2018), o Critério de Informação de Akaike avalia a qualidade do ajuste do modelo paramétrico estimado pelo método de máxima verossimelhança, assim, consiste em escolher um modelo em que as ordens k e l minimizem o critério. A equação que descreve este critério é expressa por:

$$AIC(k, d, l) = N \ln \widehat{\sigma}_a^2 + \frac{N}{N-d} 2(k+l+d+1+\delta_{d0}) + N \ln 2\pi + N \quad (3.38)$$

Onde:

$\widehat{\sigma}_a^2$ - é o estimador de máxima verossimelhança de σ_a^2 .

$$\delta_{d0} = \begin{cases} 1, & d = 0 \\ 0, & d \neq 0 \end{cases}$$

O cálculo para obter este estimador foi desenvolvido no anexo (I) deste trabalho. Com N fixado na equação, os dois últimos termos da equação foram desconsiderados, assim podemos reescrever como sendo:

$$AIC(k, l) = N \ln \widehat{\sigma}_a^2 + 2(k+l+d+2) \quad (3.39)$$

Como o objetivo é determinar as ordens p e q do modelo, então é necessário estipular valores para os limites superior K e L , já que são funções de N , ou seja, $K = L = \ln(N)$. Reescrevendo, tem-se:

$$AIC(k, l) = \ln \widehat{\sigma}_{k,l}^2 + \frac{2(k+l)}{N} \quad (3.40)$$

Segundo MORETTIN; TOLOI (2018), dependendo dos valores de K e L alguns modelos têm que ser ajustados com a finalidade de se obter o mínimo de AIC. Por exemplo, para os modelos $AR(p)$, o critério de AIC, se reduz a:

$$AIC(k) = N \ln \widehat{\sigma}_k^2 + 2K \quad (3.41)$$

SHIBATA (1976), mostra que minimizar o modelo de AIC, fornece estimativas incoerentes da ordem verdadeira do processo. Dessa forma, HURVICH; TSAI (1989) propõem uma correção para o modelo que diminui a probabilidade de selecionar uma ordem maior do que a verdadeira, assim a equação que descreve essa correção é dada por:

$$AIC_c(k) = AIC_c(k) + \frac{2(k+1)(k+2)}{N-k+2}, \quad k \leq K. \quad (3.42)$$

Após algumas simulações realizadas pelos autores foi mostrado que a correção era necessária quando N era pequeno ou quando K era uma fração grande de N . Assim, a extensão de Akaike é dada por:

$$AIC_\alpha(k) = N \ln \widehat{\sigma}_k^2 + \alpha K. \quad (3.43)$$

Onde:

α - é uma constante.

Assim, utilizando-se o critério de informação de Akaike (AIC) é possível encontrar uma métrica para comparar os valores da série e realizar a seleção do modelo. Existem outros critérios que podem ser utilizados para a seleção do melhor modelo como: Critério de Informação Bayesiano (BIC), Critério de Hannan e Quinn (HQC), Critério Final Predictor Error (FPE) e Critério Autoregressive Transfer Function (CAT).

4 Aplicação

A metodologia de séries temporais aplicada a modelos de classe paramétrica possui aplicações em diversas áreas do conhecimento. Neste capítulo foi abordada uma análise descritiva dos dados e a metodologia de séries temporais com o intuito de ajustar um modelo que melhor descreva as despesas da Secretaria da Saúde do Estado do RS e assim realizar uma projeção para os próximos 6 meses.

4.1 Software R

O software R é uma linguagem de programação que consiste numa ferramenta para cálculo e manipulação de dados. Na análise estatística o software institui o cenário para a tomada de decisão, além de permitir a plotagem de gráficos, de modo a apresentar os resultados de forma mais compreensível. Neste cenário, foi criado o RStudio, um ambiente de desenvolvimento integrado (IDE) de código aberto para criar e executar código R, que apesar de não ser necessário para trabalhar com o R, facilita ainda mais sua utilização (SCHMULLER, 2019). O software foi fundamental para este estudo, pois possibilitou realizar todas as análises de séries temporais e avaliar o melhor modelo para descrever a série. O código utilizado para as análises encontra-se nos anexos deste trabalho.

4.2 Dados

Os dados utilizados neste estudo foram obtidos do Portal Transparência do Estado do RS, disponível em : <https://www.transparencia.rs.gov.br/despesas/>, mais especificamente da Secretaria da Saúde do Estado do Rio Grande do Sul. O conjunto de dados é composto por 62 parâmetros, aos quais foram selecionados oito para compreender onde o recurso era aplicado. A Figura 5 mostra os parâmetros escolhidos e suas especificações de acordo com layout de gastos do Portal Transparência- RS.

Figura 6 – Quadro com as especificação dos parâmetros utilizados

Parâmetros	Significado	Especificação
Exercício	anual	Exercício da Fase do Gasto (Transação).
Mês	valor numérico que se refere a fração do ano de exercício.	Mês do exercício da Fase do Gasto (Transação)
Categoria	são duas: despesa de capital e despesa corrente	Classificação das receitas e despesas em operações correntes ou de capital, objetivando propiciar elementos para uma avaliação do efeito econômico das transações do setor público. (Categoria Econômica).
Subfunção	São todas as classificações contábeis a que se destinam os recursos.	A subfunção representa uma partição da função saúde, que agrega determinado subconjunto de despesas do Setor Público.
Ação	coordenadoria regional ou municipal a que se destina o recurso	Gastos que contribuem para os objetivos do programa governamental.
Projeto	é a classificação que recebe o recurso através do qual o programa atingirá seu objetivo.	Tipo de ação destinada a alcançar o objetivo de um programa.
Programa	instrumentos de ação para o cumprimento da função saúde.	Desdobramento da classificação orçamentária da despesa.
Valor	É composto por empenho, liquidação, pagamento e retenção.	Valor do Gasto de acordo com a Fase do Gastos.

Fonte: Adaptado a partir do layout de gastos disponível no Portal Transparência -RS.

Após essa compreensão preliminar, foram definidos os quatro parâmetros de interesse para o estudo (Figura 6), quais sejam: exercício, mês, categoria e valor.

Figura 7 – Planilha dos parâmetros utilizados

	A	B	C	D	E	F	G
	Categoria	Exercício	Mes	Valor			
1							
2	DESPESA CAPITAL	2018	1	2379814.63			
3	DESPESA CAPITAL	2018	2	14913787.95			
4	DESPESA CAPITAL	2018	3	4680545.35			
5	DESPESA CAPITAL	2018	4	3495989.09			
6	DESPESA CAPITAL	2018	5	16312558.5			
7	DESPESA CAPITAL	2018	6	23192475.95			
8	DESPESA CAPITAL	2018	7	6665605.88			
9	DESPESA CAPITAL	2018	8	3360886.32			
10	DESPESA CAPITAL	2018	9	6852004.03			
11	DESPESA CAPITAL	2018	10	10430067.28			
12	DESPESA CAPITAL	2018	11	8886205.6			
13	DESPESA CAPITAL	2018	12	1184313.21			
14	DESPESA CAPITAL	2019	1	1156903.64			
15	DESPESA CAPITAL	2019	2	460990.08			
16	DESPESA CAPITAL	2019	3	1038672.18			
17	DESPESA CAPITAL	2019	4	3026889.1			
18	DESPESA CAPITAL	2019	5	974717.44			
19	DESPESA CAPITAL	2019	6	86040.02			
20	DESPESA CAPITAL	2019	7	275699.34			
21	DESPESA CAPITAL	2019	8	1886085.66			
22	DESPESA CAPITAL	2019	9	6286908.27			
23	DESPESA CAPITAL	2019	10	14349087.39			
24	DESPESA CAPITAL	2019	11	2918838.91			
25	DESPESA CAPITAL	2019	12	14495622.21			
26	DESPESA CORRENTE	2018	1	985289552.06			
27	DESPESA CORRENTE	2018	2	1248644439.44			
28	DESPESA CORRENTE	2018	3	925824646.94			
29	DESPESA CORRENTE	2018	4	1063253497.36			

Fonte:Elaborado pelas autoras

4.3 Análise Descritiva

A análise descritiva inicia-se convertendo os dados brutos para escala logarítmica (log), uma vez que os dados do estudo giram na casa dos milhões e bilhões. Segundo (LÜTKEPOHL; XU, 2012), a transformação das variáveis para a escala (log) é muito utilizada em estudos relacionados a economia, pois permite uma melhor modelagem dos dados. Posteriormente, foi realizada a leitura dos dados, visando ter um panorama geral do que foi investido em despesa corrente e de capital.

Cabe destacar que a despesa corrente está relacionada com a despesas de custeio aplicadas na manutenção e no funcionamento do serviço público, por exemplo, a despesa orçamentária com pessoal. A Tabela 1 mostra os valores mensais investidos para esta despesa de janeiro de 2018 a julho de 2021. Em contrapartida, a Tabela 2 mostra parâmetros da estatística descritiva para estas despesas.

Tabela 1 – Investimento mensal em despesa corrente (log)

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2018	8.99	9.10	8.97	9.03	9.03	8.96	9.03	8.99	9.01	8.90	8.91	9.28
2019	8.81	8.95	8.90	9.04	8.98	9.00	9.04	9.03	9.06	9.11	9.05	9.29
2020	8.89	9.00	9.11	9.08	9.00	9.18	9.15	9.00	9.19	9.34	9.10	9.27
2021	8.59	8.61	8.89	7.77	9.00	9.12	9.09					

Tabela 2 – Estatística descritiva para despesa de corrente ao longo deste período (log)

	Valor
Min.	7.77
1st Qu	8.96
Median	9.02
Mean	8.99
3 rd Qu.	9.09
Max.	9.34

Observa-se que o menor valor investido ao longo deste período foi de 7.77, que ao ser convertido é de aproximadamente R\$ 5 milhões de reais. Além disso, a média é de 8.99 que corresponde a R\$ 1 bilhão de reais e o máximo investido foi de 9.34 que corresponde a R\$ 2 bilhões de reais, observa-se que o valor máximo corresponde ao mês de outubro de 2020, o que configura o cenário crítico da pandemia no Estado.

A despesa de capital, está relacionada ao investimento de capital para contribuir na geração de novos bens ou serviços que passam a integrar o serviço público, como por exemplo, compra e instalação de equipamentos hospitalares.

Tabela 3 – Investimento mensal em despesa capital (log)

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2018	6.38	7.17	6.67	6.54	7.21	7.40	6.92	6.53	6.84	7.02	6.95	6.33
2019	6.06	5.66	6.02	6.48	5.99	4.93	5.44	6.30	6.80	7.16	6.47	7.16
2020	4.10	5.00	6.73	7.29	6.31	6.37	6.47	7.49	6.86	7.67	6.93	6.98
2021	5.62	4.23	5.67	5.46	7.18	6.80	7.10					

Tabela 4 – Estatística descritiva das despesas de capital (log)

	Valor
Min.	4.09
1st Qu	6.04
Median	6.54
Mean	6.43
3 rd Qu.	7.00
Max.	7.67

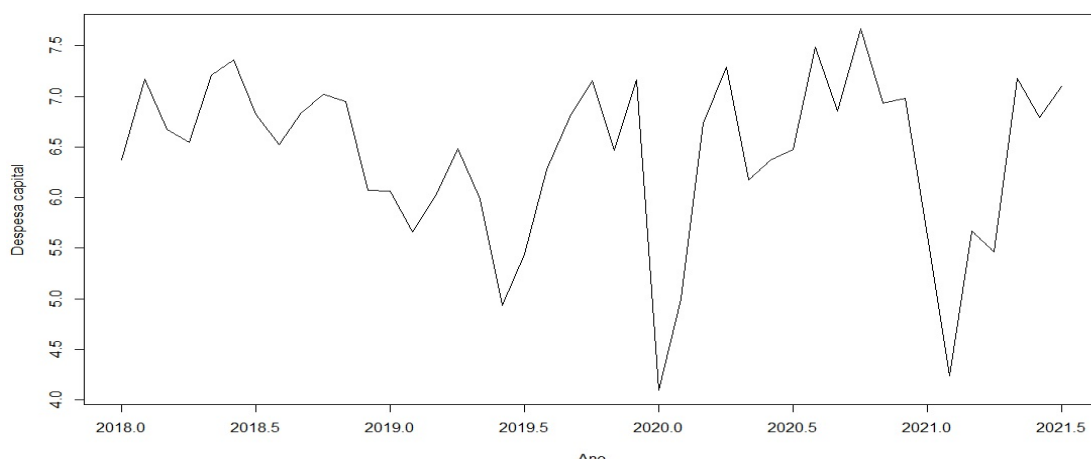
Observa-se que o menor valor mensal investido ao longo deste período foi 4.09, que ao ser convertido corresponde a R\$ 12 mil reais. A média resultou em 6.43 que convertido

é de R\$ 2 milhões de reais e o máximo investido foi cerca de 7.67 que corresponde a R\$ 47 milhões de reais, assim como as despesas correntes, a despesa de capital também atinge o valor máximo em outubro do mesmo ano.

4.4 Análise de Séries Temporais

Na análise de dados a plotagem dos gráficos permite visualizar padrões, observações incomuns, mudanças ao longo do tempo e as relações entre variáveis. Então, a partir dos dados brutos, foi realizado a conversão para série temporal utilizando a função `ts`, com intuito de descrever as características dos dados ao longo do tempo e determinar o melhor modelo para realizar previsões futuras. Deste modo, as séries que representam as despesas de capital e a despesa corrente resultaram, respectivamente, na Figura 8 e na Figura 9.

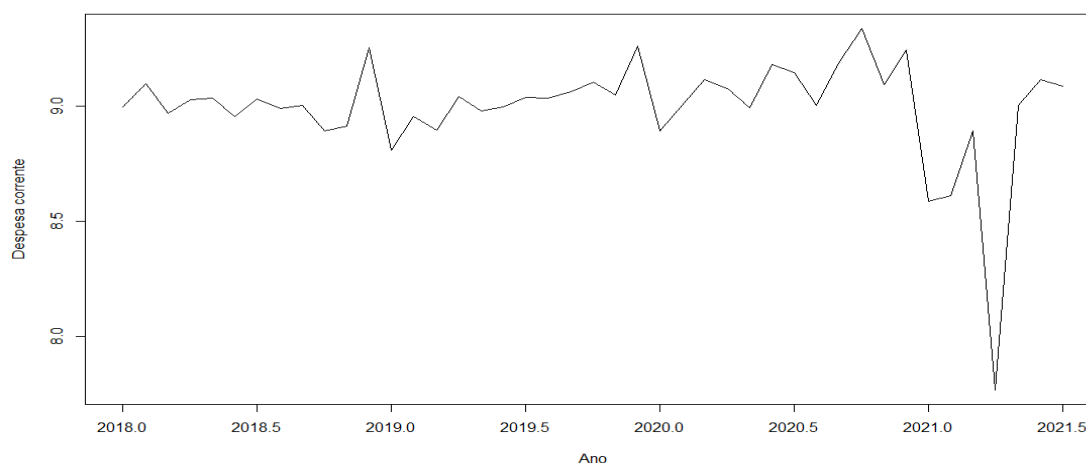
Figura 8 – Representação gráfica da série temporal para despesa de capital



Fonte: Elaborado pela autora, 2021.

Por meio da Figura 8 observa-se o crescimento a despesa de capital do primeiro para o segundo semestre de 2020. Isso ocorre porque no início da pandemia os leitos hospitalares disponibilizados para o tratamento da Covid-19 eram suficientes, mas com a disseminação do vírus houve sobrecarga nos hospitais (KREUTZ et al., 2021). Dessa forma, foi necessário a criação de hospitais de campanha para atender a população, além da aquisição de mais equipamentos hospitalares. Portanto, em decorrência do avanço da pandemia, houve demanda para o aumento da despesa de capital em saúde pública.

Figura 9 – Representação gráfica da série temporal para a despesa corrente

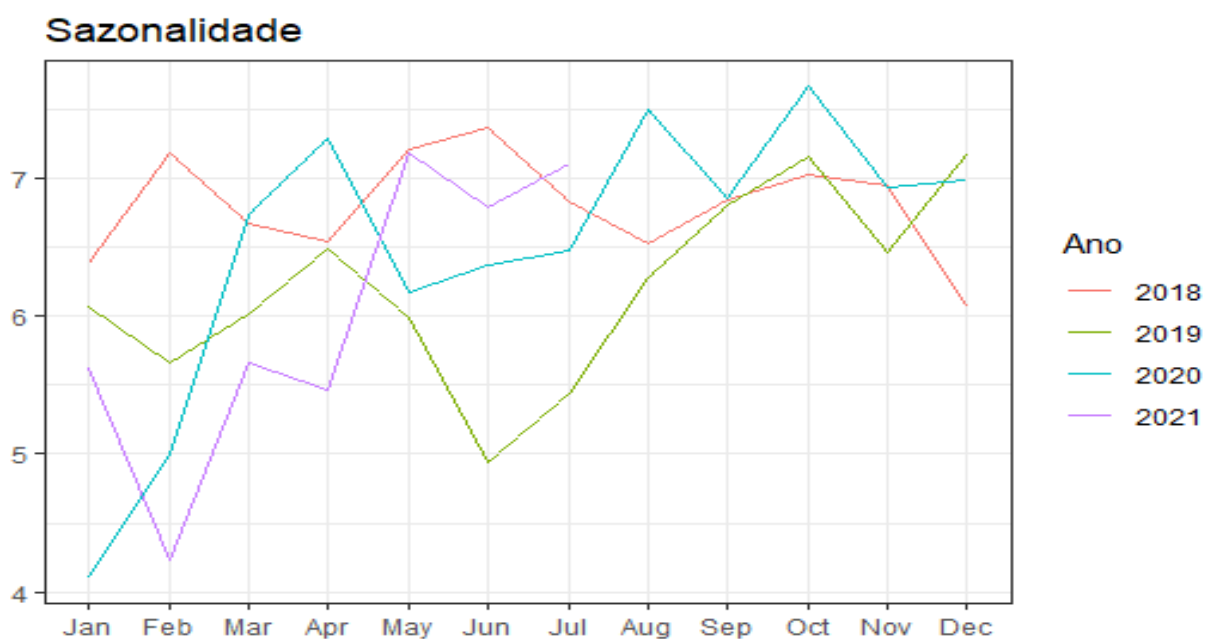


Fonte: Elaborado pela autora, 2021.

Da mesma forma, observa-se um aumento da despesa corrente no segundo semestre do ano de 2020, o que também decorre do crescimento de casos positivos de Covid-19 no RS. Consequentemente, foi necessário investir na contratação de pessoal e na aquisição de EPI'S, assim como na manutenção de equipamentos para o tratamento dos infectados pelo vírus (KREUTZ et al., 2021). Apesar disso, no primeiro semestre de 2021, é possível visualizar a queda nesses investimentos, pois com o início da vacinação, o número de internações diminuiu, logo, não foi mais necessário contratação em massa de pessoal e os equipamentos de UTI'S disponíveis já eram suficientes para atender a demanda.

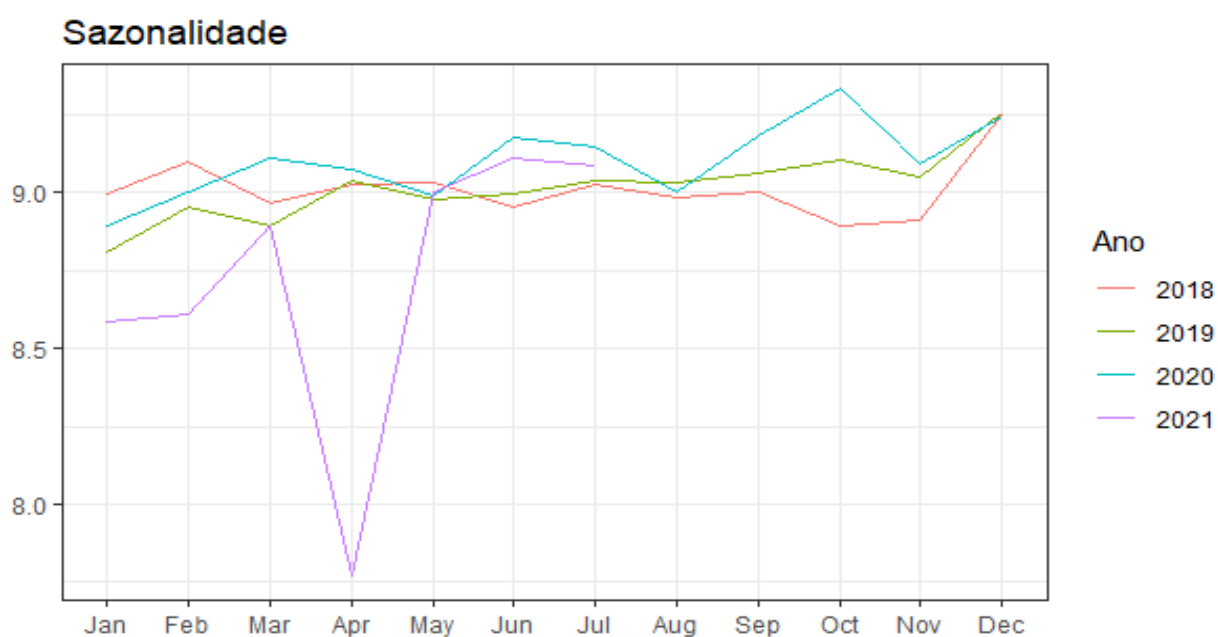
Analisando os gráficos das séries de despesas é possível ver que o comportamento da série não possui oscilações regulares (periódicas) ao longo do tempo, não podendo ser classificada como sazonal. A Figura 10 e 11 mostra a sazonalidade para as despesas ao longo do tempo. Assim, como as oscilações da série se repetem ao longo do tempo de forma irregular a série pode ser classificada como uma série cíclica (HYNDMAN, et al., 2002).

Figura 10 – Sazonalidade das despesas de capital



Fonte: Elaborado pela autora.

Figura 11 – Sazonalidade das despesas corrente



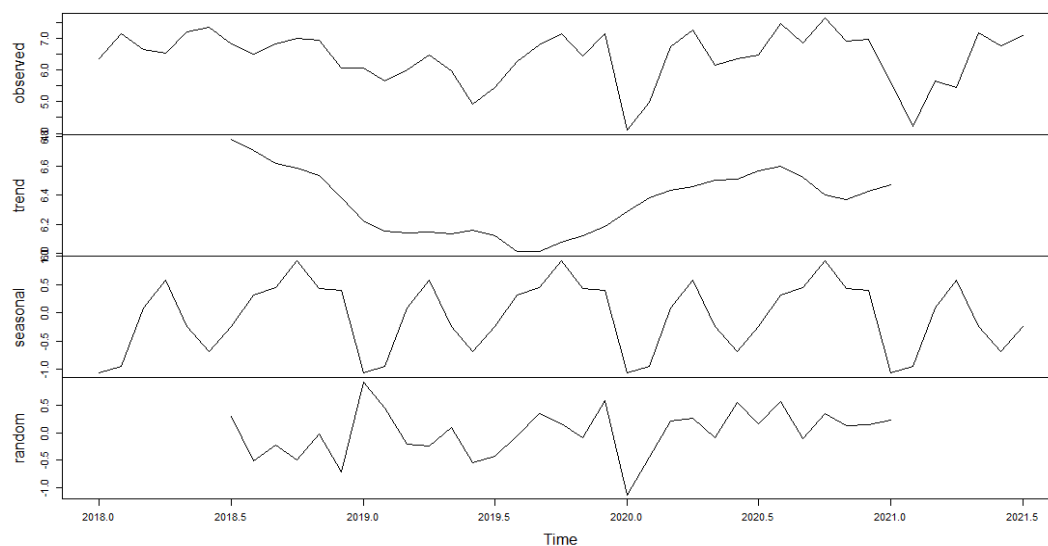
Fonte: Elaborado pela autora.

4.4.1 Decomposição

Para descrever o comportamento da série é necessário observar cada uma de suas componentes, pois através delas é possível classificar a série e assim identificar o melhor ajuste. Desta forma, as Figuras 12 e 13 representam, respectivamente, a decomposição

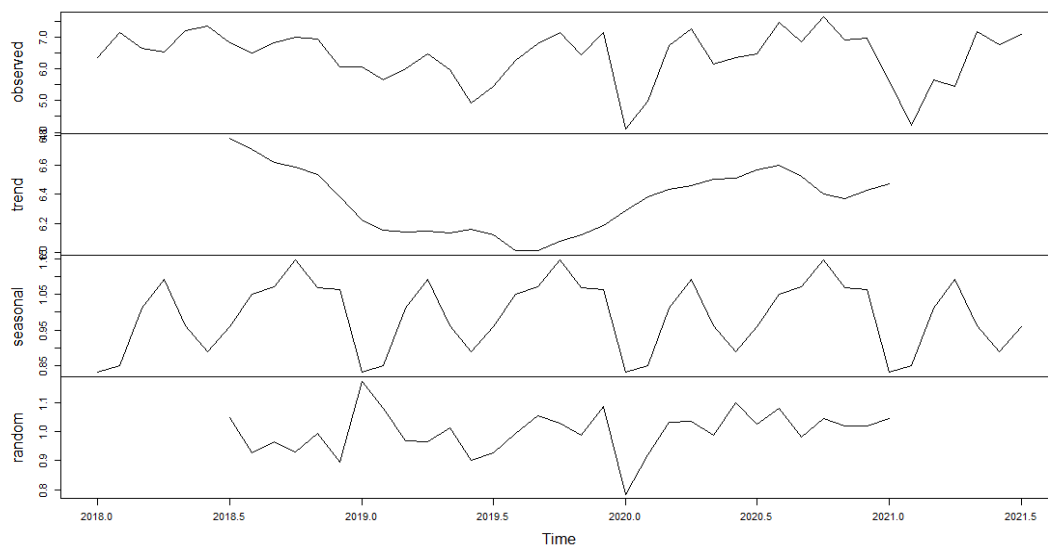
aditiva e multiplicativa para a série da despesa de capital.

Figura 12 – Decomposição aditiva para a série de despesa de capital



Fonte: Elaborado pela autora.

Figura 13 – Decomposição multiplicativa para a série de despesa de capital



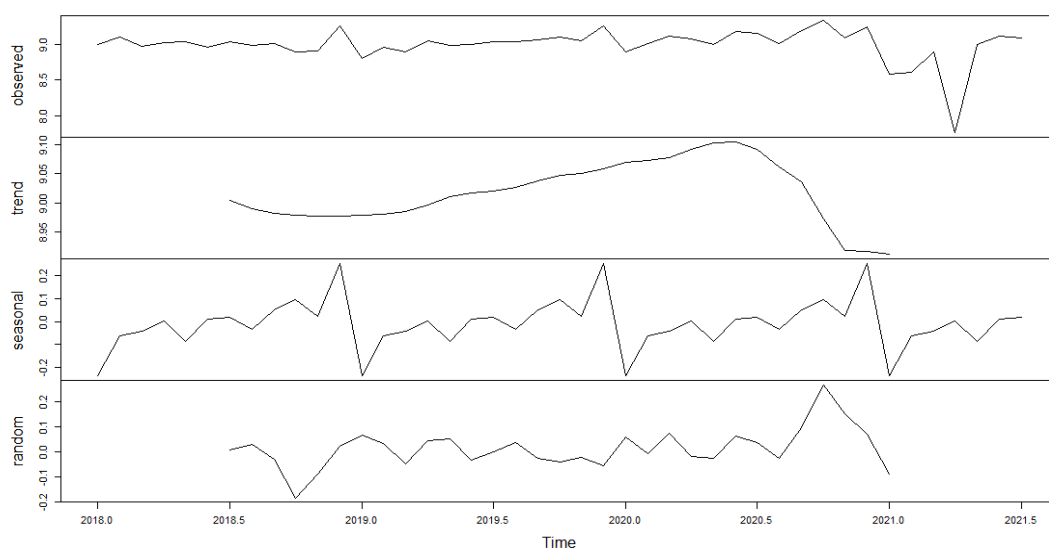
Fonte: Elaborado pela autora, 2021.

É importante mencionar que a decomposição é composta por 3 componentes: tendência, sazonalidade e resíduo. Cada um desses componentes descreve o comportamento da série. Além disso, a observação descreve a série temporal original dos dados, isto é, todas as observações.

- A tendência mostra a mudança de direção, por exemplo, quando a série passa por mudança de crescente para decrescente.
- A sazonalidade, por sua vez, indica a frequência fixa conhecida.
- O resíduo é uma série de variáveis aleatórias não correlacionadas, cuja expectativa é zero e sua variância é constante.

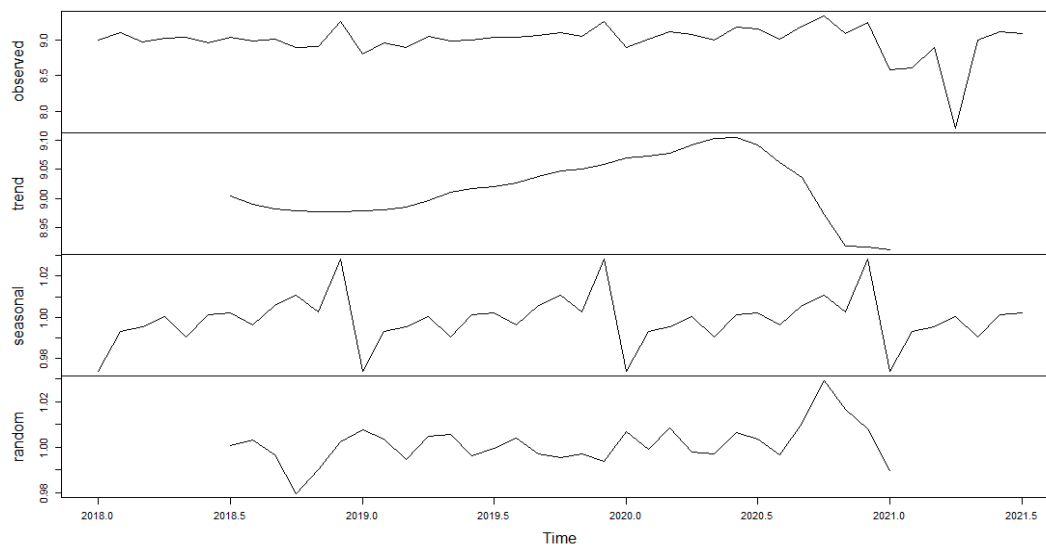
Já a decomposição aditiva e multiplicativa para a série de despesa corrente está ilustrada nas Figuras 14 e 15.

Figura 14 – Decomposição aditiva para a série de despesa de corrente



Fonte: Elaborado pela própria autora, 2021.

Figura 15 – Decomposição multiplicativa para a série de despesa corrente



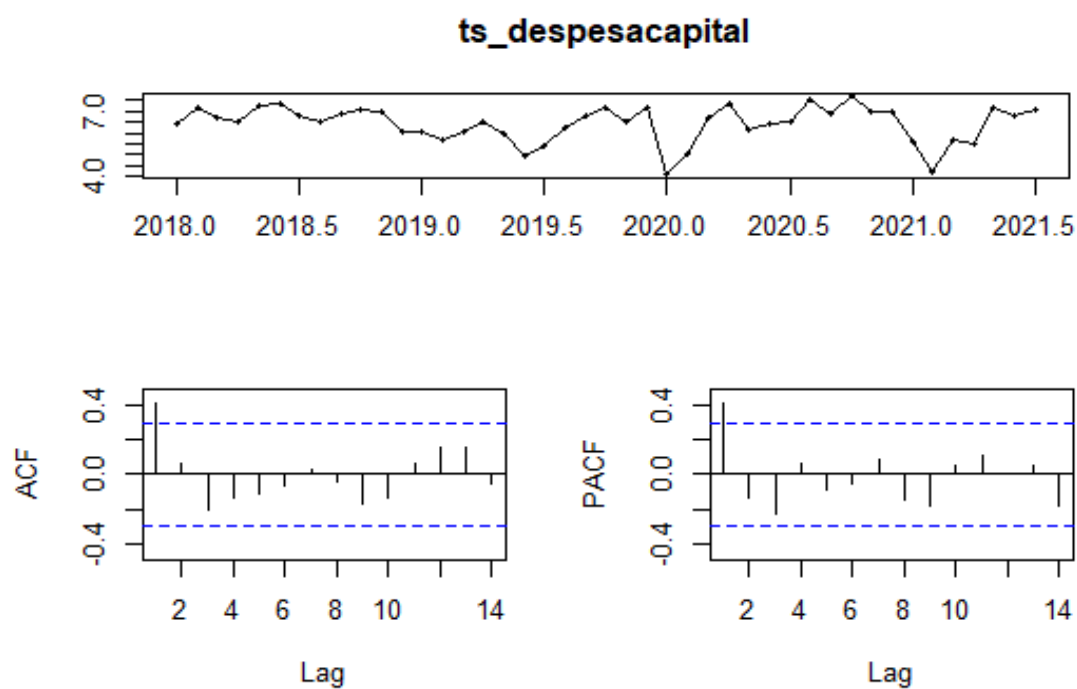
Fonte: Elaborado pela própria autora, 2021.

4.4.2 Estacionariedade

A estacionariedade de uma série temporal pode ser avaliada através das funções de autocorrelação parcial, ou ainda, pelo uso de alguns testes tais como: testes ADF, PP e kPSS descritos na seção 3.12.

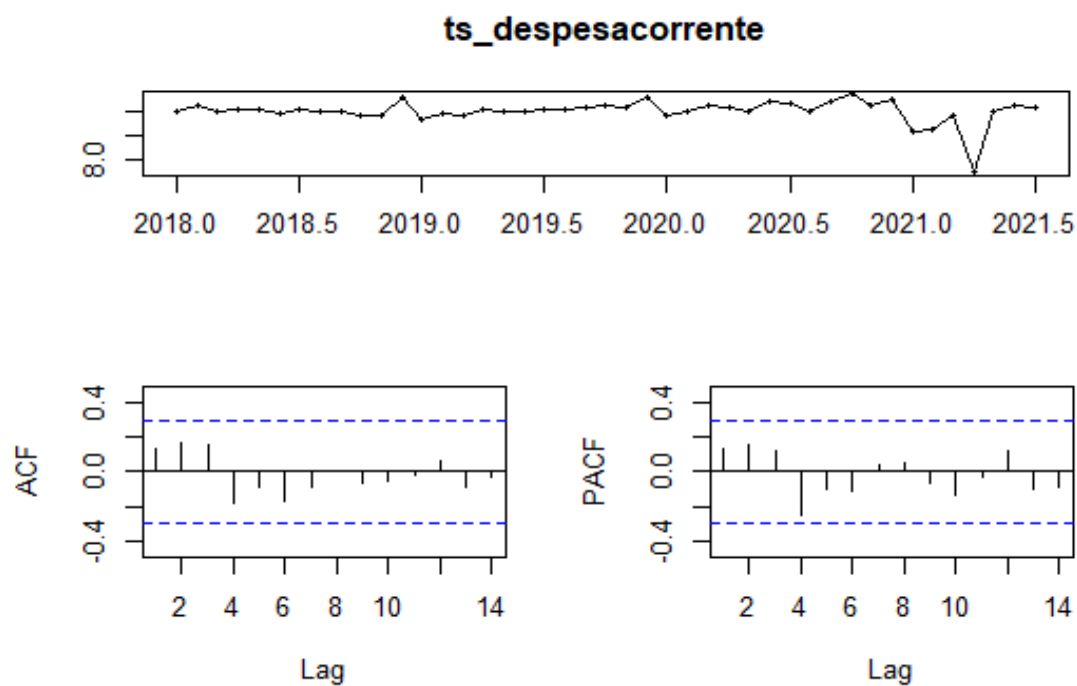
Assim, avaliando a estacionariedade através da função de autocorrelação parcial (PACF), obtem-se os resultados expressos pelas Figuras 16 e 17. Em relação a autocorrelação (ACF), observa-se que os valores estão dentro do limite estabelecido pela linha azul, significando que a autocorrelação é zero. Portanto, é possível classificar a série como aleatória.

Figura 16 – Função de autocorrelação para despesa de capital



Fonte: Elaborado pelas autoras, 2022.

Figura 17 – Função de autocorrelação para despesa corrente



Fonte: Elaborado pelas autoras, 2022.

Também foram analisados a autocorrelação dos resíduos por meio do teste Box-Ljung, os resultados são expressos pela Tabela 5.

Tabela 5 – Teste Box-Ljung

X-squared	df	lag	p-valor
6.4923	0	1	< 2.2e-16
8.0018	1	2	0.004673
8.8015	2	3	0.01227
8.8619	3	4	0.03118

Analisando os resultados da tabela, a primeira coluna x-squared indica a estatística do teste para verificar se os resíduos são independentes, os df indicam os graus de liberdade, lag é o período de tempo que separa os dados e o p-valor indica a probabilidade que mede a evidência contra a hipótese nula.

Para o teste de normalidade foi utilizado Shapiro-Wilk que apresentou um $W=0.92083$ e $p\text{-valor} = 0.005713$. Esses resultados mostram que não há rejeição da hipótese nula.

Além disso, avaliando a estacionariedade através dos testes mencionados na seção 3.12 deste trabalho obtém-se os resultados elencados nas Tabelas 6 e 7.

Tabela 6 – Estatística dos testes para despesa de capital

Teste	Valor	p-valor
ADF	-3.0056	0.1765
PP	-23.5740	0.01537
KPPS	0.0795	0.1000

Tabela 7 – Estatística dos testes para despesa corrente

Teste	Valor	p-valor
ADF	-4.1608	0.1251
PP	-40.5270	0.01000
KPPS	0.0963	0.1000

De acordo com os testes aplicados, tem-se que parte dos testes os valores calculados encontram-se dentro do limite estabelecido pelo p-valor que para os testes de ADF e PP é de $p < 0.05$ e para o teste KPPS é de $p > 0.05$ de significância, analisando os testes aplicados conclui-se que para o ADF o $p > 0.05$ o que rejeita-se a hipótese nula, já para o testes PP e KPPS o p-valor encontra-se dentro dos limites estabelecidos, então pode-se concluir que a série é estacionária. Portanto, conhecendo as características da série foi possível realizar o ajuste de modelos para descrever seu comportamento e realizar previsões.

4.4.3 Ajuste do modelo

Dessa forma, verificada a estacionariedade, iniciou-se o ajuste da série aos modelos. O primeiro ajuste a ser feito foi o da sazonalidade para tentar encontrar uma periodicidade

na série. Os resultados para este ajuste podem ser verificados através do teste QS que avalia a qualidade do ajuste sazonal.

Tabela 8 – Teste de ajuste sazonal para despesa de capital

	Série	qs	p-val
Original		1.16	0.56
original corrigida por valores extremos		14.99	0.00
Resíduos do modelo de regressão		1.15	0.56
Série com ajuste sazonal		0.00	1.00
Série sem ajuste sazonal		0.00	1.00
Componente irregular		0.00	1.00
Componente irregular corrigido com valores extremos		0.00	1.00

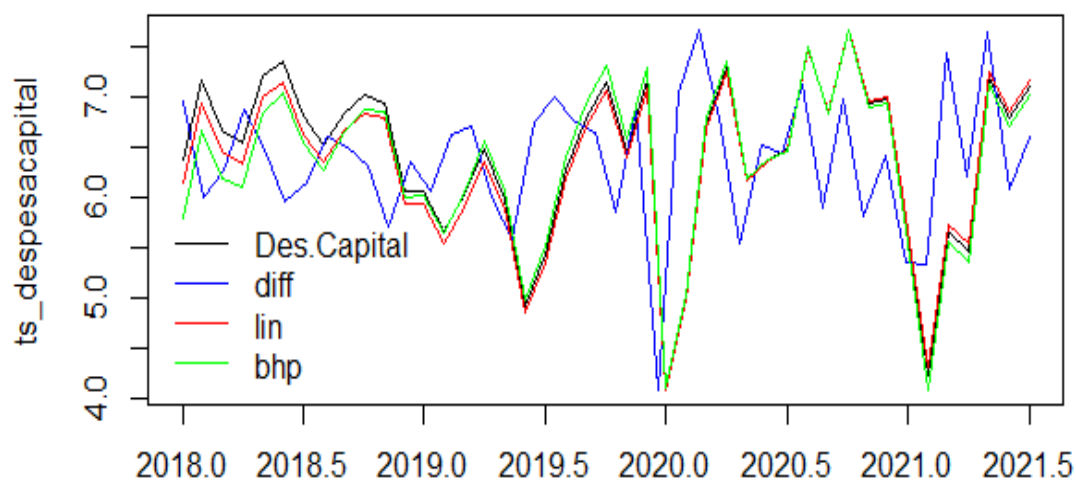
Tabela 9 – Teste de sazonalidade para despesa corrente

	Série	qs	p-val
Original		1.16	0.56
original corrigida por valores extremos		14.99	0.00
Resíduos do modelo de regressão		1.15	0.56
Série com ajuste sazonal		0.00	1.00
Série sem ajuste sazonal		0.00	1.00
Componente irregular		0.00	1.00
Componente irregular corrigido com valores extremos		0.00	1.00

Por meio dos dados expressos nas Tabelas 8 e 9, é possível observar que o ajuste sazonal não é necessário neste caso, pois a série não sofre alteração quando se remove a componente sazonal.

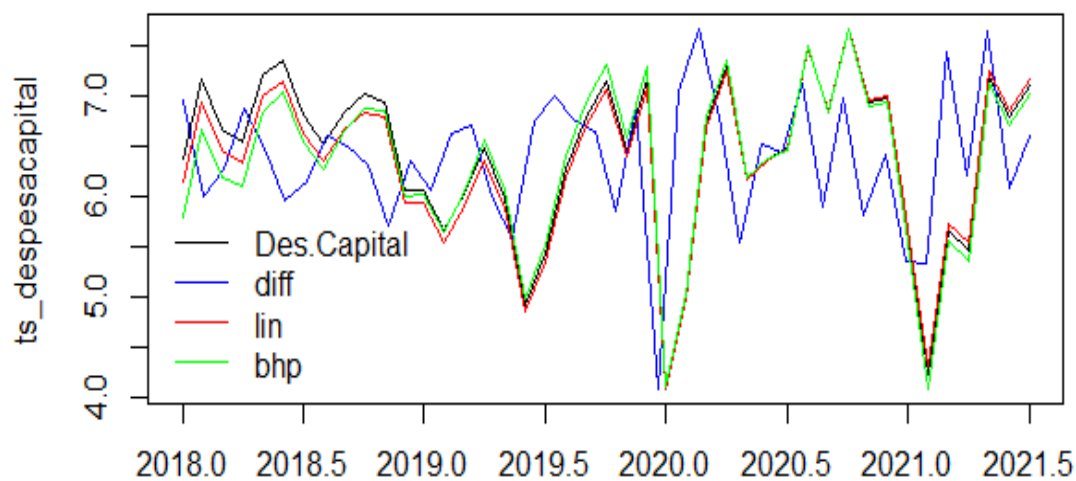
Adicionalmente, tentou-se ajustar a componente tendência através de três métodos: primeira diferença da série (diff), remoção da tendência linear (lin), remoção da tendência pelo filtro HODRICK-PRESCOTT (bhp), o qual tem por objetivo remover as flutuações que possam ocorrer na série. Os resultados dos três filtros foram representados nas Figuras 18 e 19.

Figura 18 – Remoção de Tendência para despesa de capital



Fonte: Elaborado pela própria autora, 2021.

Figura 19 – Remoção de Tendência para despesa corrente



Fonte: Elaborado pela própria autora, 2021.

Os gráficos mostram que os diferentes filtros aplicados apresentam o mesmo com-

portamento da série, logo, o ajuste de tendência torna-se insignificante.

Após, o teste de tendência e sazonalidade foi realizado o ajuste através do modelo ARIMA, conforme os resultados mostrados nas Tabelas 10 e 11. O melhor ajuste do modelo para despesa de capital foi obtido por um ARIMA de ordem (2,0,1), com um AIC=101.78 sendo uma generalização do modelo ARMA(2,1). Destaca-se que o AIC quanto menor, melhor o ajuste para a série.

Tabela 10 – Ajuste ARIMA para despesa de capital

Modelo ARIMA	ϕ_1	ϕ_2	θ_1	θ_2	Erro	AIC
(1,0,1)	0.2964			0.1548	6.4417	103.86
(1,0,2)	-0.3178		0.8261	0.4635	6.4318	103.35
(2,0,1)	1.3375	-0.5034	-1.0000		6.3861	101.78
(2,0,2)	-0.2889	-0.2714	0.7778	0.6296	6.4298	104.46
(1,1,1)	0.4506		-1.0000			103.67
(1,1,2)	0.3454			-0.8665	-0.1335	105.37
(2,1,1)	0.4975	-0.1182	-1.0000			105.90
(2,2,2)	0.5548	-0.0833	-1.9895	0.9998		113.35

Assim, a equação que descreve o modelo dado pela ordem (2,0,1) é :

$$\widehat{Y}_t = \mu + 1.3375Y_{t-1} - 0.5034Y_{t-2} + 1.0000e_{t-1} \quad (4.1)$$

Conforme a Tabela 10, o melhor ajuste para despesa corrente se dá por um modelo ARIMA(2,0,2) com um AIC =2.25.

Tabela 11 – Ajuste ARIMA para despesa de corrente

Modelo ARIMA	ϕ_1	ϕ_2	θ_1	θ_2	Erro	AIC
(1,0,1)	0.5448			-0.4003		5.48
(1,0,2)	0.3678		-0.3340	0.3893	9.0059	5.68
(2,0,1)	0.2633	0.1348	-0.1592		8.9977	6.95
(2,0,2)	0.5067	-0.4586	-0.5218	1.0000	9.0070	2.25
(1,1,1)	0.1549		-1.0000			8.55
(1,1,2)	-0.9084		0.0617	-0.9383		10.45
(2,1,1)	0.1368	0.1739	-1.0000			9.26
(2,2,2)	-1.2114	-0.5867	-0.3715	-0.6285		19.38

Assim, a equação que descreve o modelo dado ARIMA(2,0,2) é :

$$\widehat{Y}_t = \mu - 0.5067Y_{t-1} - 0.4586Y_{t-2} + 0.5218e_{t-1} - 1.0000e_{t-2} \quad (4.2)$$

E por fim foi feito o ajuste dos dados através do modelo TAR (Autoregressive Threshold Model) tem por objetivo aproximar séries não-lineares, em séries lineares por partes. Recordando que o modelo é dado por TAR(p,d,q), onde p é a ordem do primeiro regime, q a ordem do segundo regime e d é a defasagem.

Para a série de despesa de capital ajustada pelo modelo TAR, de ordem (2,0,1), com AIC=59.88. Assim pode-se dizer que estes três modelos fazem um ajuste dos dados em dois regimes e sem desfasagem.

Tabela 12 – Ajuste TAR para despesa de capital

Modelo TAR	AIC
(1,0,1)	60.45
(1,0,2)	61.23
(2,0,1)	59.88
(2,0,2)	61.79
(1,1,1)	98.02
(1,1,2)	98.28
(2,1,1)	94.19
(2,2,2)	95.45

Assim, o modelo TAR(2,0,1) para esta despesa poder ser expresso através da equação:

$$y_t = \begin{cases} 7.45 - 0.1353y_{t-1} - 0.2510y_{t-2} + \epsilon_t & se \quad y_{t-1} \leq 5.66 \\ 5.71 + 0.1614y_{t-1} + \epsilon_t & se \quad y_{t-1} > 5.66 \end{cases} \quad (4.3)$$

Em contrapartida, para as séries de despesa corrente o ajuste foi TAR(1,0,1) com um AIC = -45.23.

Tabela 13 – Ajuste TAR para despesa corrente

Modelo TAR	AIC
(1,0,1)	-45.23
(1,0,2)	-41.11
(2,0,1)	-41.10
(2,0,2)	-39.11
(1,1,1)	-28.62
(1,1,2)	-24.58
(2,1,1)	-25.80
(2,2,2)	-13.32

A equação a seguir descreve o ajuste para o modelo TAR(1,0,1).

$$y_t = \begin{cases} 7.83 + 0.0551y_{t-1} + \epsilon_t & se \quad y_{t-1} \leq 8.61 \\ 8.44 + 0.0637y_{t-1} + \epsilon_t & se \quad y_{t-1} > 8.61 \end{cases} \quad (4.4)$$

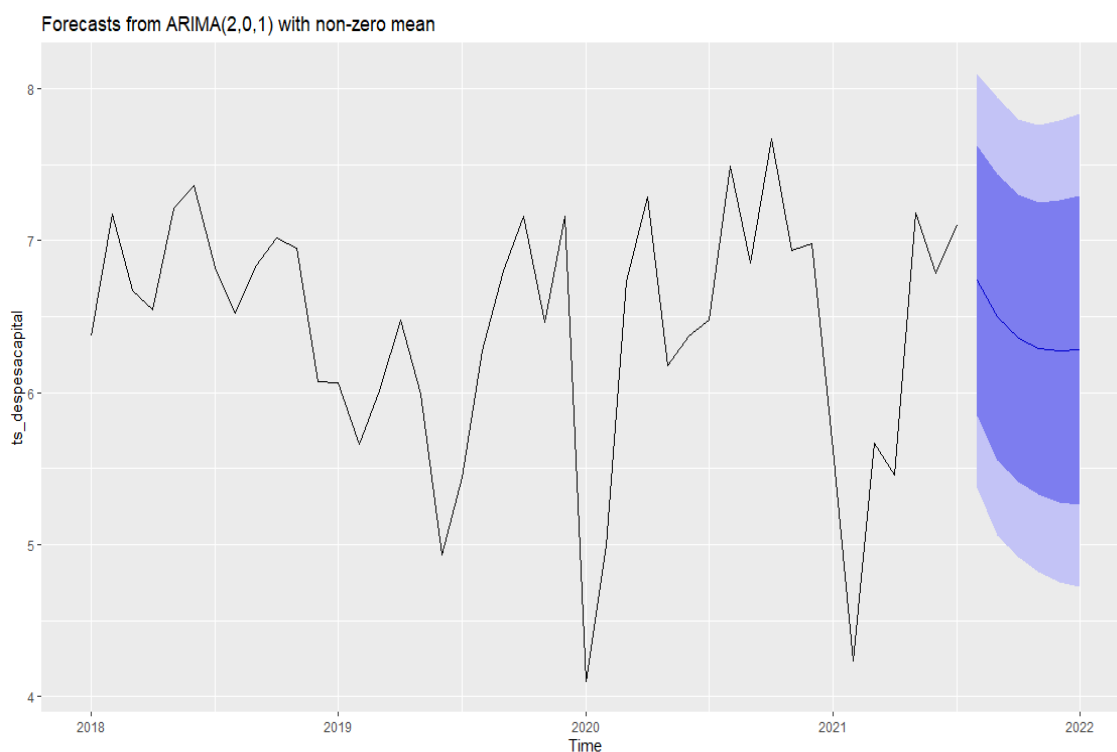
4.4.4 Previsão

Para a previsão optou-se por realizar uma projeção primeiramente através dos modelos ARIMA, foco principal deste trabalho, para estimar quanto será investido nos próximos 6 meses. A Tabela 14 mostra a projeção para o modelo ARIMA, melhor ajustado, para a despesa de capital e a Figura 20 representa a projeção do modelo graficamente.

Tabela 14 – Previsão pelo modelo ARIMA (2,0,1) para a despesa de capital

Mês	Point Forecast	Lo 80	Hi 80	Lo 95	Hi95
Agosto(2021)	6.737461	5.846970	7.627952	5.375572	8.099349
Setembro (2021)	6.498008	5.556191	7.439826	5.057622	7.938394
Outubro (2021)	6.358625	5.416560	7.300690	4.917860	7.799389
Novembro (2021)	6.292116	5.330482	7.253751	4.821423	7.762810
Dezembro (2021)	6.272775	5.281449	7.264100	4.756673	7.788876
Janeiro (2022)	6.279963	5.263726	7.296200	4.725763	7.834163

Figura 20 – Previsão ARIMA para despesa de capital



Fonte: Elaborado pela própria autora, 2022.

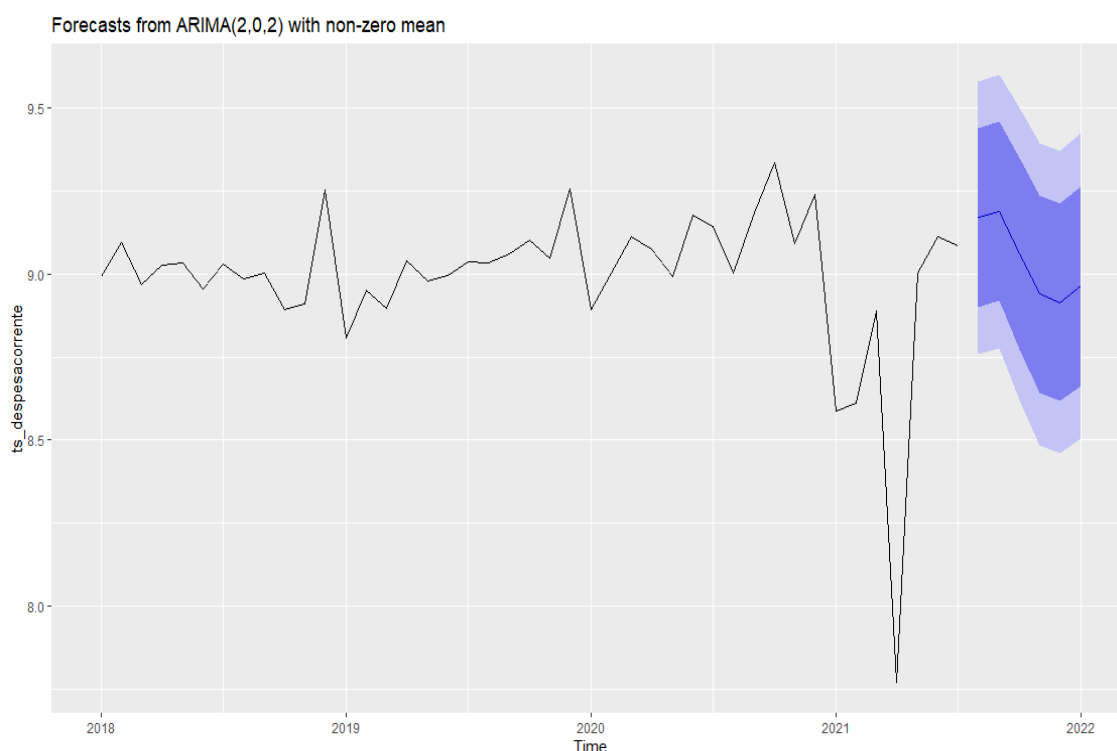
Observando-se a Figura 20 em relação as despesas de capital, conclui-se que o valor total ao longo do período não irá ultrapassar 7.0, o que corresponde a investimento em torno de R\$ 10 milhões de reais. Na Figura 20, as linhas pretas representam a série em estudo , a linha azul as previsões para os próximos 6 meses fora da série em estudo, a região sombreada escura mostra o intervalo de previsão de 80% e a região sobreada clara mostra o intervalo de 95 % .

Para a despesa corrente a melhor ordem foi o ARIMA (2,0,2), as previsões são mostradas na Tabela 15 e a Figura 21 representa graficamente esses valores.

Tabela 15 – Previsão pelo modelo ARIMA (2,0,2) para a despesa de corrente

Mês	Point Forecast	Lo 80	Hi 80	Lo 95	Hi95
Agosto(2021)	9.168937	8.899883	9.437991	8.757455	9.580419
Setembro (2021)	9.188106	8.918705	9.457508	8.776093	9.600120
Outubro (2021)	9.059988	8.772597	9.347379	8.620461	9.499515
Novembro (2021)	8.939503	8.642091	9.236915	8.484651	9.394355
Dezembro (2021)	8.913940	8.616492	9.211387	8.459032	9.368847
Janeiro (2022)	8.964531	8.664178	9.264885	8.505180	9.423883

Figura 21 – Previsão ARIMA para despesa corrente



Fonte: Elaborado pela própria autora, 2022.

Já para as despesas correntes observa-se que o total não irá ultrapassar 9.10. Assim, estima-se que o investimento seja em torno de R\$ 1,2 bilhões de reais. Na Figura 21, as linhas pretas representam a série em estudo , a linha azul as previsões para os próximos 6

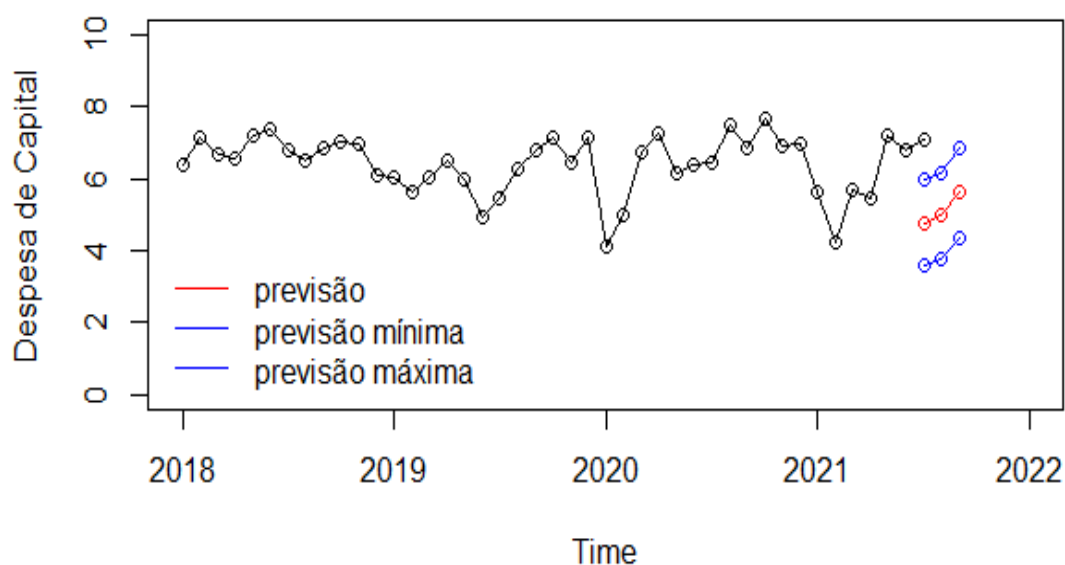
meses fora da série em estudo, a região sombreada escura mostra o intervalo de previsão de 80% e a região sobreada clara mostra o intervalo de 95 % .

Em um segundo momento fez-se a projeção para os 6 meses futuros, das despesas de capital através do modelo TAR melhor ajustado, valores descritos na Tabela 16 e pela Figura 21.

Tabela 16 – Ajuste da despesa de capital

	prevmin	prevmax	previsao
Agosto (2021)	3.57	6.03	4.80
Setembro (2021)	3.79	6.28	5.03
Outubro (2021)	4.28	6.82	5.55
Novembro (2021)	4.18	6.68	5.42
Dezembro (2021)	4.06	6.60	5.33
Janeiro (2021)	4.09	6.64	5.37

Figura 22 – Representação gráfica do modelo TAR para despesa de capital



Fonte: Elaborado pela própria autora, 2021.

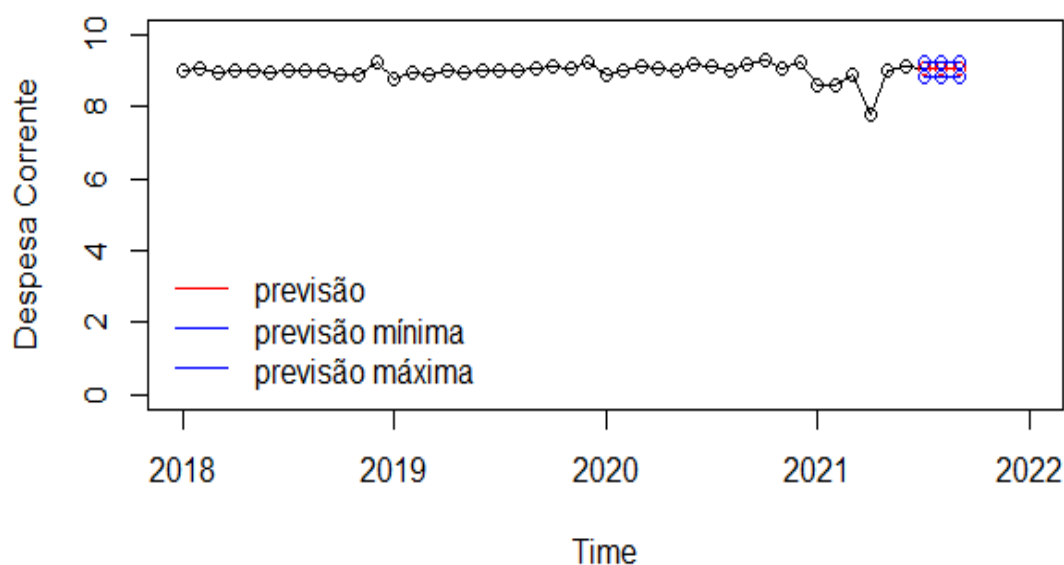
Observa-se que os investimentos, em despesas de capital, por essa projeção, não irá ultrapassar 5.55 o que corresponde cerca de R\$ 354 mil.

Em relação a despesa corrente os investimentos não ultrapassam 9.03 que é em torno de 1 bilhão de reais.

Tabela 17 – Ajuste TAR para despesa corrente

	prevmin	prevmax	previsao
Agosto (2021)	8.72	9.35	9.03
Setembro (2021)	8.71	9.34	9.02
Outubro (2021)	7.60	9.40	9.00
Novembro (2021)	7.60	9.41	8.98
Dezembro (2021)	7.57	9.73	8.98
Janeiro (2022)	7.47	9.75	8.97

Figura 23 – Representação gráfica do modelo TAR para despesa corrente



Fonte: Elaborado pela própria autora, 2021.

Os modelos ARIMA são mais consistentes na estimativa de valores futuros que os modelos TAR, que por vezes superestimam (despesas de capital) e por outras subestimam os valores (despesas corrente). Percebe-se que as previsões são bem conservadoras em ambos os casos, sem grandes oscilações, mostrando a dificuldade de se prever os valores futuros de um conjunto de dados com características tão peculiares: despesas de um Estado frente a uma pandemia inesperada.

5 Considerações Finais

O presente estudo permitiu avaliar os investimentos realizados na secretaria da saúde do estado do Rio Grande do Sul, no período do primeiro semestre de 2018 ao primeiro semestre de 2021, utilizando a metodologia de séries temporais. Observou-se que os dados apresentam ausência de sazonalidade e tendência e apresentam um comportamento estacionário.

Preliminarmente, a estatística descritiva permitiu a compreensão dos dados para melhor entendimento de suas características e dimensão.

O modelo ARIMA com melhor ajuste foi ARIMA(2,0,1) para despesas de capital e ARIMA(2,0,2) para despesas corrente. Dentre os modelos TAR estimados, o melhor ajuste para despesas de capital foi pelo modelos TAR(2,0,1) e para as despesas correntes por um modelo TAR(1,0,1). O melhor ajuste dos modelos foi determinado com base no critério de Akaike.

Em se tratando das previsões dos valores para os 6 meses a frente dos dados coletados, os modelos ARIMA se mostraram mais consistentes na previsão dos valores futuros que os modelos TAR, que por vezes superestimaram os valores (despesas de capital) e por outras subestimaram os dados (despesas corrente).

Percebe-se que as previsões são bem conservadoras em ambos os casos, sem grandes oscilações, mostrando a dificuldade de se prever os valores futuros de um conjunto de dados. Deste modo, foi possível, no decorrer deste trabalho, analisar de forma descritiva as despesas da Secretaria da Saúde do RS, antes e durante pandemia de Covid-19, bem como realizar projeções de investimentos públicos na área para os próximos 6 meses com base nos dados passados.

Referências

- ACHUTTI, C. F. Análise do planejamento orçamentário da secretaria estadual de saúde do estado do rio grande do sul (ses/rs): uma visão a partir dos créditos adicionais suplementares. 2019.
- BOSQUE, L. M. Análise de séries temporais não-lineares com modelos auto-regressivos threshold. 2014.
- BOX, G. E.; JENKINS, G. M. Time series analysis. forecasting and control. *Holden-Day Series in Time Series Analysis*, 1976.
- BRASIL, G. F. *SUS completa 30 anos da criação*. 2020. Disponível em: <<https://www.gov.br/casacivil/pt-br/assuntos/noticias/2020/setembro/sus-completa-30-anos-da-criacao>>.
- CAVALARO, N. L. Modelo de séries temporais autorregressivo periódico-par. Universidade Federal de São Carlos, 2021.
- CHEIN, F. *Introdução aos modelos de regressão linear*. Brasília-DF: [s.n.], 2019.
- COSTA, N. S. S. d. Detecção de quebras estruturais em séries temporais: Implementação dos testes de shimotsu com uma aplicação em séries do mercado de câmbio. 2016.
- CUNHA, D. *Administração Pública: uma visão ampla da administração pública direta e indireta*. 2015. 2016.
- FRACARO, N. Estacionariedade das séries temporais do modelo matemático arimax de propulsores eletromecânicos. 2018.
- GIOVANELLA, L. et al. *Políticas e sistema de saúde no Brasil*. [S.l.]: SciELO-Editora FIOCRUZ, 2012.
- GLOROT, X.; BENGIO, Y. Understanding the difficulty of training deep feedforward neural networks. In: JMLR WORKSHOP AND CONFERENCE PROCEEDINGS. *Proceedings of the thirteenth international conference on artificial intelligence and statistics*. [S.l.], 2010. p. 249–256.
- GOMES, F. Modelo adaptativo para predição de nível de líquidos em cadinhos de altos-fornos baseado em séries temporais. 2016.
- HURVICH, C. M.; TSAI, C.-L. Regression and time series model selection in small samples. *Biometrika*, Oxford University Press, v. 76, n. 2, p. 297–307, 1989.
- HYNDMAN, R. J.; ATHANASOPOULOS, G. *Forecasting: Principles and Practice*. second. Austrália: Texts, 2018. v. 2.
- KEILA. *Séries Temporais*. 2015. Disponível em: <https://www.maxwell.vrac.puc-rio.br/24787/24787_3.PDF>.

- KREUTZ, I. M. et al. As coordenadorias regionais de saúde e o enfrentamento à covid-19: A experiência da 14ª coordenadoria regional de saúde, rio grande do sul (rs), brasil. In: *Congresso Internacional em Saúde*. [S.l.: s.n.], 2021.
- LAMOUNIER, W. M. Tendência, ciclos e sazonalidade nos preços spot do café brasileiro na nybot. *Gestão & Produção*, SciELO Brasil, v. 14, n. 1, p. 13–23, 2007.
- LIMA, E. A. d. et al. Modelagem de séries temporais em dados de temperatura do ar por semivariografia univariada: interpolação spline e validação cruzada. Universidade Federal de Mato Grosso, 2015.
- LÜTKEPOHL, H.; XU, F. The role of the log transformation in forecasting economic variables. *Empirical Economics*, Springer, v. 42, n. 3, p. 619–638, 2012.
- MORETTIN, P. A.; TOLOI, C. M. C. *Análise de Séries Temporais*. third. São Paulo: Blucher, 2018. v. 1.
- OLIVEIRA, M. F. D. Metodologia científica: um manual para a realização de pesquisas em administração. *Universidade Federal de Goiás. Catalão-GO*, 2011.
- OLIVEIRA, U.; CARVALHO, F. d. S. Comparação de técnicas de previsão de demanda para controle de estoques de embalagem para computadores. *ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO*, v. 33, p. 1–18, 2013.
- PEDROSA, D. P. F. *Ajuste de Curvas*. 2021. Disponível em: <<https://www.dca.ufrn.br/~diogo/FTP/dca0304/ajustedecurvas.pdf>>.
- REIS, M. M. *ANÁLISE DE SÉRIES TEMPORAIS*. 2020.
- RELVAS, T. R. S. Previsões com séries temporais: cálculo de coeficientes sazonais com regressão linear. In: *Anais do Congresso Brasileiro de Custos-ABC*. [S.l.: s.n.], 1998.
- SANTOS, N. M. dos. A organização da saúde no rio grande do sul. *Revista Gestão & Saúde*, v. 1, n. 3, sep 2010.
- SANTOS, P. D. A. B. *ESTATÍSTICA: definição e conceitos básicos*. 2019. Disponível em: <<https://www.ibilce.unesp.br/Home/Departamentos/CiencCompEstatistica/Adriana/conceitos-basicos---estatistica.pdf>>.
- SANTOS, R. de Cássia Leal Fonseca dos. *Orçamento Público*. third. [S.l.]: rev. atual., 2016.
- SCHMULLER, J. *Análise Estatística com R para leigos*. second. Rio de Janeiro: Alta Books, 2019. v. 1.
- SHIBATA, R. Selection of the order of an autoregressive model by akaike's information criterion. *Biometrika*, Oxford University Press, v. 63, n. 1, p. 117–126, 1976.
- SILVEIRA, A. G.; MATTOS, V. L. D. de; KONRATH, A. C. Avaliação da estacionariedade e teste de cointegração em séries temporais o caso da demanda de energia elétrica residencial no brasil. *RETEC-Revista de Tecnologias*, v. 9, n. 3, 2016.
- SOEIRO, J. de F. et al. Modelos de séries temporais aplicados ao consumo de energia elétrica do estado do amapá (pp. 127). *Revista da Estatística da Universidade Federal de Ouro Preto*, v. 2, 2012.

TRANSPARÊNCIA, P. da. *Portal da Transparência-RS*. 2021. Disponível em: <<http://portaltransparencia.gov.br/entenda-a-gestao-publica/execucao-despesa-publica-despesapublica>>.

TRIOLA, M. F. *Introdução à Estatística*. tenth. [S.l.: s.n.], 2017.

Apêndice A- Estimador de Máxima Verossimilhança

O método de máxima verossimilhança consiste em encontrar os estimadores $\hat{\theta}_t = \hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k$ que maximizem o logaritmo da função de verossimilhança dos dados (BATISTA, 2015). Assim, seja a função dada por:

$$L(\theta_1, \theta_2, \dots, \theta_K) = \sum_{i=1}^n y_i(\theta_1 + \theta_2 x_1 + \dots + \theta_k x_k) - m_i \ln(1 + e^{\theta_1 + \theta_2 x_1 + \dots + \theta_k x_k}) \quad (5.1)$$

Derivando a equação em relação a cada um dos estimadores, obtem-se:

$$\frac{\partial L}{\partial \theta_1} = \sum_{i=1}^n y_i - m_i \ln \frac{e^{\theta_1 + \theta_2 x_1 + \dots + \theta_k x_k}}{1 + e^{\theta_1 + \theta_2 x_1 + \dots + \theta_k x_k}} \quad (5.2)$$

$$\frac{\partial L}{\partial \theta_2} = \sum_{i=1}^n y_i x_i - m_i \ln \frac{e^{\theta_1 + \theta_2 x_1 + \dots + \theta_k x_k}}{1 + e^{\theta_1 + \theta_2 x_1 + \dots + \theta_k x_k}} \quad (5.3)$$

$$\frac{\partial L}{\partial \theta_k} = \sum_{i=1}^n y_i x_k - m_i \ln \frac{e^{\theta_1 + \theta_2 x_1 + \dots + \theta_k x_k}}{1 + e^{\theta_1 + \theta_2 x_1 + \dots + \theta_k x_k}} \quad (5.4)$$

Igualando cada uma das derivadas a zero:

$$\frac{\partial L}{\partial \theta_1} = \sum_{i=1}^n y_i - m_i \ln \frac{e^{\theta_1 + \theta_2 x_1 + \dots + \theta_k x_k}}{1 + e^{\theta_1 + \theta_2 x_1 + \dots + \theta_k x_k}} = 0 \quad (5.5)$$

$$\frac{\partial L}{\partial \theta_2} = \sum_{i=1}^n y_i x_i - m_i \ln \frac{e^{\theta_1 + \theta_2 x_1 + \dots + \theta_k x_k}}{1 + e^{\theta_1 + \theta_2 x_1 + \dots + \theta_k x_k}} = 0 \quad (5.6)$$

$$\frac{\partial L}{\partial \theta_k} = \sum_{i=1}^n y_i x_k - m_i \ln \frac{e^{\theta_1 + \theta_2 x_1 + \dots + \theta_k x_k}}{1 + e^{\theta_1 + \theta_2 x_1 + \dots + \theta_k x_k}} = 0 \quad (5.7)$$

E encontrando as equações para cada uma das derivadas acima, obtem-se os valores para os estimadores, logo basta retornar a equação inicial e encontrar o valor máximo da função.

Apêndice B - Código das Análises

```

1 library(dplyr)
2 library(tidyverse)
3 library(forecast)
4 library(tseries)
5 library(Amelia)
6 library(xtable)
7 library(windows)
8 #LEITURA DOS DADOS
9 dadospc<-read.table("despesaporcategoria.csv",dec = ".",sep=",",header =
    TRUE)
10 dadosdca<-read.table("despesadecapital.csv",dec = ".",sep=",",header = TRUE
    )
11 dadosdco<-read.table("despesacorrente.csv",dec = ".",sep=",",header = TRUE)
12 dadosfinais<-rbind(dadospc , dadosdca , dadosdco)
13 str(dadosfinais)
14
15 #IDEIA GERAL DOS DADOS
16 summary(dadosdca$Valor)
17 log10(dadosdca$Valor)
18 summary(log10(dadosdca$Valor))
19 summary(dadosdco$Valor)
20 log10(dadosdco$Valor)
21 summary(log10(dadosdco$Valor))
22
23 #OGARITMO DOS DADOS BRUTOS
24 #TRANSFORMANDO PARA SERIES TEMPORAIS
25 ts_despesacapital=ts(data=log10(dadosdca$Valor),start = c(2018,1),end=c
    (2021,7),frequency = 12)
26 ts_despesacapital
27
28 ts_despesacorrente=ts(data=log10(dadosdco$Valor),start = c(2018,1),end=c
    (2021,7),frequency = 12)
29 ts_despesacorrente
30
31 #OBSERVACAO DE UM PERIODO ESPECIFICO DA SERIE UTILIZAMOS "HEAD" PARA O
    INICIO E "TAIL" PARA O FIM
32 head(ts_despesacapital,n=6)
33 head(ts_despesacorrente,n=6)
34 tail(ts_despesacapital,n=6)
35 tail(ts_despesacorrente,n=6)
36
37 #POSICAO DE CADA OBSERVACAO
38 cycle(ts_despesacapital)

```

```
39 cycle(ts_despesacorrente)
40
41 plot(ts_despesacapital, xlab = "Ano", ylab = "Despesa capital")
42 plot(ts_despesacorrente, xlab = "Ano", ylab = "Despesa corrente")
43
44 #HISTOGRAMA
45 hist(ts_despesacapital)
46 hist(ts_despesacorrente)
47
48 #DECOMPONDO A SERIE
49 x.decompClassic<-decompose(ts_despesacapital, type = "additive")
50 plot(x.decompClassic)
51 x.decompClassic<-decompose(ts_despesacorrente, type = "additive")
52 plot(x.decompClassic)
53
54 x.decompLoess<-decompose(ts_despesacapital, type="multiplicative")
55 plot(x.decompLoess)
56 x.decompLoess<-decompose(ts_despesacorrente, type="multiplicative")
57 plot(x.decompLoess)
58
59 #ESTACIONARIDADE
60
61 library(tseries)
62 library(urca)
63 library(htest)
64
65 summary(ur.df(ts_despesacapital, type='none', lags=0))
66 adf.test(ts_despesacapital)
67 xtable(adf.test(ts_despesacapital))
68 pp.test(ts_despesacapital)
69 kpss.test(ts_despesacapital, null = "Trend")
70 adf.test(ts_despesacorrente)
71 pp.test(ts_despesacorrente)
72 kpss.test(ts_despesacorrente, null = "Trend")
73
74 #AUTOCORRELAÇÃO
75
76 tsdisplay(ts_despesacapital)
77 tsdisplay(ts_despesacorrente)
78 ggtsdisplay(ts_despesacapital)
79 ggtsdisplay(ts_despesacorrente)
80 Acf(ts_despesacapital)
81 Acf(ts_despesacorrente)
82
83 #AUTOCORRELAÇÃO PARCIAL
84 Pacf(ts_despesacapital)
85 Pacf(ts_despesacorrente)
```

```

86
87 library(ggplot2)
88
89 ggseasonplot(ts_despesacapital) +
90   ggtitle("Sazonalidade") +
91   labs(x = "") +
92   scale_colour_discrete(name = "Ano") +
93   theme_bw()
94
95 monthplot(ts_despesacapital, col='darkblue', col.base='red',
96           main='Despesa por mês e Despesa mensal',
97           ylab='Despesa de Capital', lty.base=2)
98
99 legend('bottomright', legend=c('Despesa de Capital', 'Mês'), lty=c
100      (1,2),
101      col=c('darkblue', 'red'), bty='n')
102
103 fit<-arma(linCapital)
104
105 fit
106
107 library(lmtest)
108 library(zoo)
109 library(lomb)
110
111 #AJUSTE ARIMA – DCORRENTE
112
113 mod.ARIMA.ca = list()
114 mod.ARIMA.ca[["Dcorrente_ARIMA01"]] = arima(ts_despesacorrente, order=c
115      (1,0,1))
116 mod.ARIMA.ca[["Dcorrente_ARIMA02"]] = arima(ts_despesacorrente, order=c
117      (1,0,2))
118 mod.ARIMA.ca[["Dcorrente_ARIMA03"]] = arima(ts_despesacorrente, order=c
119      (2,0,1))
120 mod.ARIMA.ca[["Dcorrente_ARIMA04"]] = arima(ts_despesacorrente, order=c
121      (2,0,2))
122 mod.ARIMA.ca[["Dcorrente_ARIMA06"]] = arima(ts_despesacorrente, order=c
123      (1,1,1))
124 mod.ARIMA.ca[["Dcorrente_ARIMA07"]] = arima(ts_despesacorrente, order=c
125      (1,1,2))
126 mod.ARIMA.ca[["Dcorrente_ARIMA08"]] = arima(ts_despesacorrente, order=c
127      (2,1,1))
128 mod.ARIMA.ca[["Dcorrente_ARIMA09"]] = arima(ts_despesacorrente, order=c
129      (2,2,2))
130 mod.ARIMA.ca
131 summary(mod.ARIMA.ca[["Dcorrente_ARIMA09"]])

```

```

124
125 tab<-sapply(mod.ARIMA.ca,AIC)
126 xtable(tab)
127
128 mod.ARIMA.ca = list()
129 mod.ARIMA.ca[["Dcorrente_ARIMA01"]] = arima(ts_despesacapital, order=c
      (1,0,0))
130
131 fit <-mod.ARIMA.ca[["Dcorrente_ARIMA03"]]
132 checkresiduals(fit)
133
134 forecast::forecast(fit,h=6)
135 autoplot(forecast::forecast(fit,h=6))
136
137 # Ajuste modelo TAR -DCAPITAL
138
139 mod.TAR.ca = list()
140 mod.TAR.ca[["Dcapital_TAR01"]] = TSA::tar(ts_despesacapital,p1=1,d=0, p2=1,
      method="MAIC")
141 mod.TAR.ca[["Dcapital_TAR01"]]$AIC
142 mod.TAR.ca[["Dcapital_TAR02"]] = TSA::tar(ts_despesacapital,p1=1,d=0, p2=2,
      method="MAIC")
143 mod.TAR.ca[["Dcapital_TAR02"]]$AIC
144 mod.TAR.ca[["Dcapital_TAR03"]] = TSA::tar(ts_despesacapital,p1=2,d=0, p2=1,
      method="MAIC")
145 mod.TAR.ca[["Dcapital_TAR03"]]$AIC
146 mod.TAR.ca[["Dcapital_TAR04"]] = TSA::tar(ts_despesacapital,p1=2,d=2, p2=2,
      method="MAIC")
147 mod.TAR.ca[["Dcapital_TAR04"]]$AIC
148 mod.TAR.ca[["Dcapital_TAR06"]] = TSA::tar(ts_despesacapital,p1=1,d=1, p2=1,
      method="MAIC")
149 mod.TAR.ca[["Dcapital_TAR06"]]$AIC
150 mod.TAR.ca[["Dcapital_TAR07"]] = TSA::tar(ts_despesacapital,p1=1,d=1, p2=2,
      method="MAIC")
151 mod.TAR.ca[["Dcapital_TAR07"]]$AIC
152 mod.TAR.ca[["Dcapital_TAR08"]] = TSA::tar(ts_despesacapital,p1=2,d=2, p2=1,
      method="MAIC")
153 mod.TAR.ca[["Dcapital_TAR08"]]$AIC
154 mod.TAR.ca[["Dcapital_TAR09"]] = TSA::tar(ts_despesacapital,p1=2,d=2, p2=2,
      method="MAIC")
155 mod.TAR.ca[["Dcapital_TAR09"]]$AIC
156 mod.TAR.ca
157 summary(mod.TAR.ca[["Dcapital_TAR09"]])
158 mod.TAR.ca[["Dcapital_TAR09"]]$AIC
159
160 #PREVISAO
161 n_prev = 6

```

```
162 tar_prev = predict(mod.TAR.ca[["Dcapital_TAR01"]], n.ahead=n_prev, n.sim =
      1000)
163 #MANIPULANDO OS DADOS DA PREVIS O
164 prev1 = t(tar_prev$pred.interval)
165 prev2 = tar_prev$fit
166 prev3 = cbind(prev1, prev2)
167 prev = as.data.frame(prev3)
168 colnames(prev)[1] = "prevmin"
169 colnames(prev)[2] = "prevmax"
170 colnames(prev)[3] = "previsao"
171
172 library(xtable)
173 library(astsa)
174 xtable(prev)
175 dadosfinais<-prev[,3]
176 dadosfinaismin<-prev[,1]
177 dadosfinaismax<-prev[,2]
178 ts_dadosfinaisprev=ts(data=dadosfinais, start = c(2021,7), end=c(2021,9),
      frequency = 12)
179 ts_dadosfinaismin=ts(data=dadosfinaismin, start = c(2021,7), end=c(2021,9),
      frequency = 12)
180 ts_dadosfinaismax=ts(data=dadosfinaismax, start = c(2021,7), end=c(2021,9),
      frequency = 12)
181 plot(ts_despesacapital, type='o', ylab='Despesa de Capital', xlim=c(2018,2022)
      ,ylim=c(0,9))
182 lines(ts_dadosfinaisprev, col="red")
183 points(ts_dadosfinaisprev, col="red")
184 lines(ts_dadosfinaismin, col="blue")
185 points(ts_dadosfinaismin, col="blue")
186 lines(ts_dadosfinaismax, col="blue")
187 points(ts_dadosfinaismax, col="blue")
188
189 #Manipulando os dados da previs o
190 prev1 = t(tar_prev$pred.interval)
191 prev2 = tar_prev$fit
192 prev3 = cbind(prev1, prev2)
193 prev = as.data.frame(prev3)
194 colnames(prev)[1] = "prevmin"
195 colnames(prev)[2] = "prevmax"
196 colnames(prev)[3] = "previsao"
197
198 xtable(prev)
199 prev
200 xdadosfinais<-prev[,3]
201 dadosfinaismin<-prev[,1]
202 dadosfinaismax<-prev[,2]
203
```

```

204 ts_dadosfinaisprev=ts(data=dadosfinais , start = c(2021,1) , end=c(2021,7) ,
    frequency = 12)
205 ts_dadosfinaismin=ts(data=dadosfinaismin , start = c(2021,1) , end=c(2021,7) ,
    frequency = 12)
206 ts_dadosfinaismax=ts(data=dadosfinaismax , start = c(2021,1) , end=c(2021,7) ,
    frequency = 12)
207
208 plot(xprev , type='o' , ylab='Despesa de Capital' , xlim=c(2018,2022) , ylim=c(0,8)
    )
209 points(ts_dadosfinaisprev , col="red")
210 points(ts_dadosfinaismin , col="blue")
211 points(ts_dadosfinaismax , col="blue")
212
213
214 # Ajuste modelo TAR-DCORRENTE
215
216 ###TAR
217 mod.TAR.co = list()
218 mod.TAR.co[["Dcorrente_TAR01"]] = TSA::tar(ts_despesacorrente , p1=1,d=0, p2
    =1,method="MAIC")
219 mod.TAR.co[["Dcorrente_TAR01"]] $AIC
220 mod.TAR.co[["Dcorrente_TAR02"]] = TSA::tar(ts_despesacorrente , p1=1,d=0, p2
    =2,method="MAIC")
221 mod.TAR.co[["Dcorrente_TAR02"]] $AIC
222 mod.TAR.co[["Dcorrente_TAR03"]] = TSA::tar(ts_despesacorrente , p1=2,d=0, p2
    =1,method="MAIC")
223 mod.TAR.co[["Dcorrente_TAR03"]] $AIC
224 mod.TAR.co[["Dcorrente_TAR04"]] = TSA::tar(ts_despesacorrente , p1=2,d=2, p2
    =2,method="MAIC")
225 mod.TAR.co[["Dcorrente_TAR04"]] $AIC
226 mod.TAR.co[["Dcorrente_TAR06"]] = TSA::tar(ts_despesacorrente , p1=1,d=1, p2
    =1,method="MAIC")
227 mod.TAR.co[["Dcorrente_TAR06"]] $AIC
228 mod.TAR.co[["Dcorrentel_TAR07"]] = TSA::tar(ts_despesacorrente , p1=1,d=1, p2
    =2,method="MAIC")
229 mod.TAR.co[["Dcorrentel_TAR07"]] $AIC
230 mod.TAR.co[["Dcorrente_TAR08"]] = TSA::tar(ts_despesacorrente , p1=2,d=2, p2
    =1,method="MAIC")
231 mod.TAR.co[["Dcorrente_TAR08"]] $AIC
232 mod.TAR.co[["Dcorrente_TAR09"]] = TSA::tar(ts_despesacorrente , p1=2,d=2, p2
    =2,method="MAIC")
233 mod.TAR.co[["Dcorrente_TAR09"]] $AIC
234 mod.TAR.co
235
236 summary(mod.TAR.co[["Dcorrente_TAR01"]])
237
238 mod.TAR.co[["Dcorrente_TAR09"]] $AIC

```

```
239
240 #Previsão
241 n_prev = 6
242
243 tar_prev = predict(mod.TAR.co[["Dcorrente_TAR01"]], n.ahead=n_prev, n.sim =
      1000)
244
245 #Manipulando os dados da previsão
246 prev1 = t(tar_prev$pred.interval)
247 prev2 = tar_prev$fit
248 prev3 = cbind(prev1, prev2)
249 prev = as.data.frame(prev3)
250 colnames(prev)[1] = "prevmin"
251 colnames(prev)[2] = "prevmax"
252 colnames(prev)[3] = "previsao"
253
254 library(xtable)
255 library(astsa)
256 prev
257 xtable(prev)
258 dadosfinais<-prev[,3]
259 dadosfinaismin<-prev[,1]
260 dadosfinaismax<-prev[,2]
261
262 ts_dadosfinaisprev=ts(data=dadosfinais, start = c(2021,7), end=c(2021,9),
      frequency = 12)
263 ts_dadosfinaismin=ts(data=dadosfinaismin, start = c(2021,7), end=c(2021,9),
      frequency = 12)
264 ts_dadosfinaismax=ts(data=dadosfinaismax, start = c(2021,7), end=c(2021,9),
      frequency = 12)
265
266 plot(ts_despesacorrente, type='o', ylab='Despesa Corrente', xlim=c(2018,2022),
      ylim=c(0,9))
267 lines(ts_dadosfinaisprev, col="red")
268 points(ts_dadosfinaisprev, col="red")
269 lines(ts_dadosfinaismin, col="blue")
270 points(ts_dadosfinaismin, col="blue")
271 lines(ts_dadosfinaismax, col="blue")
272 points(ts_dadosfinaismax, col="blue")
```
