

UM ESTUDO DE ANÁLISE E PREVISÃO DE SÉRIES TEMPORAIS APLICADO ÀS CURVAS DE INDICES DE ISOLAMENTOS DEVIDOS À PANDEMIA DE COVID-19

Diego Clementino Pompeo e Arnaldo Rabello de Aguiar Vallim Filho (Orientador)

Apoio: PIBIC Mackenzie

RESUMO

Nesta pesquisa, foi feita uma análise e previsão das séries temporais do índice de isolamento devido a pandemia da COVID-19, durante um período de um ano, fevereiro de 2020 a fevereiro de 2021, das principais cidades das 16 regiões administrativas do estado de São Paulo, divididas em 4 sub-regiões: a Capital, a Região Metropolitana de São Paulo (RMSP), o Litoral e o Interior. Para que fosse possível a análise e previsão do índice de isolamento de cada uma dessas regiões, foram utilizados os dados disponibilizados pelo governo de Estado de São Paulo, que após a limpeza, a transformação e a integração, foi possível o desenvolvimento de quatro tipos de análise: cálculo de média móvel, curva de tendência e curva de sazonalidade (anual e semanal), e ainda a previsão para períodos futuros. As previsões foram feitas para até duas semanas à frente e com uma margem de erro estipulada dentro no próprio algoritmo desenvolvido. Todo o estudo foi desenvolvido com o apoio da linguagem de programação R, e suas bibliotecas, tornando assim, possível a compreensão e análise do comportamento dos índices de isolamento de cada município analisado.

Palavras-chave: Séries Temporais, COVID-19, Ciência de Dados

ABSTRACT

In this research, an analysis and forecast of the time series of the isolation index due to the COVID-19 pandemic was carried out, over a period of one year, from February 2020 to February 2021, in the main cities of the 16 administrative regions of the state of São Paulo, divided into 4 sub-regions: the Capital, the Metropolitan Region of São Paulo (RMSP), the Coast and the Interior. In order to analyze and predict the isolation index of each of these regions, data provided by the São Paulo State government were used, which after cleaning, transformation and integration, it was possible to develop four types of analysis: calculation of moving average, trend curve and seasonality curve (annual and weekly), as well as forecasting for future periods. The predictions were made for up to two weeks ahead and with a margin of error stipulated within the developed algorithm itself. The entire study was developed with the support of the R programming language, and its libraries, thus making it possible to understand and analyze the behavior of the isolation indices of each analyzed municipality.

Keywords: Time Series, COVID-19, Data Science

1. INTRODUÇÃO

A doença causada pelo coronavírus, a COVID-19, teve seu primeiro caso registrado na China no final de 2019, e em março de 2020, a OMS declarou o surto como pandemia. Após essa declaração da OMS, diversos países adotaram o lockdown (confinamento) como forma de prevenção contra a COVID-19, visto que, por hora, não se tem vacinas e nem medicamentos 100% eficazes contra a doença.

Em alguns países, mesmo adotando o confinamento como forma de prevenção, o número de pessoas contaminadas ainda continua crescente, devido a população não respeitar o confinamento, causando assim, um número crescente também de mortes, como é o caso do Brasil (Observatório COVID-19 BR, 2020)

A evolução da pandemia vem se desenvolvendo de forma acelerada no país, conforme mostram os dados oficiais do Ministério da Saúde e das Secretarias de Saúde dos estados da federação (Ministério da Saúde, 2020).

A previsão do que deve ocorrer nos próximos períodos, considerando-se a evolução das séries temporais de casos registrados, poderia representar um apoio relevante, de forma a propiciar um adequado planejamento das ações dos administradores públicos e privados. E, uma vez que tenham modelos eficazes de análise e previsão, estes poderiam se constituir em ferramentas de apoio a estudos sobre outros tipos de pandemia, que podem vir a ocorrer.

Inúmeros estudos estão sendo desenvolvidos em todo o mundo, e há uma quantidade grande de dados disponíveis, o que é um ponto favorável para o desenvolvimento de pesquisas associadas à Ciência de Dados, que é o caso desta proposta de projeto, que poderá assim, ajudar a compreender melhor esse fenômeno, especialmente pelas técnicas de Previsão de Séries Temporais, que permitem analisar um conjunto de observações ordenadas no tempo, e prever seu comportamento futuro.

Assim, é que surge esta proposta de projeto de Iniciação Científica. No caso específico deste projeto, o trabalho contemplou a análise e previsão de séries temporais de variáveis relevantes associadas à pandemia de COVID-19, estudando séries de tempo de dados internos do país e séries de outros países, particularmente, as séries cronológicas de índices de isolamento das pessoas nos municípios do estado de São Paulo.

A Estatística, a Inteligência Artificial (IA) e a Ciência de Dados, de forma geral, oferecem um conjunto de técnicas, algoritmos e ferramentas que permitiriam o desenvolvimento desse tipo de estudo. A questão que se coloca é quais seriam os instrumentos mais adequados para determinados conjuntos de dados.

Assim, pensando-se em formular e caracterizar de forma bem delimitada, um projeto de pesquisa voltado para esse tema, a pergunta de pesquisa para a qual se pretende buscar uma resposta, seria:

Conhecendo-se as técnicas estatísticas clássicas de análise e previsão de séries temporais e algoritmos de IA, voltados para o mesmo objetivo, como se comportariam as previsões desenvolvidas quando aplicadas a séries de tempo do índice de isolamento de pessoas, devidos à pandemia de COVID-19, em diferentes regiões geográficas do estado de São Paulo?

O Objetivo Geral da pesquisa, assim, foi desenvolver modelos de análise e previsão de séries temporais com o uso de algoritmos utilizados em Ciência de Dados.

Em termos de Objetivos Específicos buscou-se desenvolver modelos de análise e previsão de séries temporais do índice de isolamento de pessoas, durante o período de pandemia, em cidades relevantes das principais regiões administrativas do estado de São Paulo. As análises deveriam ser desenvolvidas com o apoio de ferramentas computacionais amplamente utilizadas em Ciência de Dados. Em termos de análise deveriam ser desenvolvidos métodos de suavização dos dados originais, como é o caso de médias móveis, curvas de tendência, curvas de sazonalidade, previsões para períodos futuros e limites de confiança para essas previsões (margens de erro).

2. REFERENCIAL TEÓRICO

Esse capítulo tem como objetivo apresentar os trabalhos científicos encontrados que possuem relação com o tema do projeto. Os materiais estudados auxiliaram no entendimento do problema e no conhecimento do tema.

Os assuntos aqui pesquisados se concentraram na pandemia, propriamente dita, e na técnica de série temporal a ser empregada no estudo. Sobre estas, procurou-se levantar na literatura trabalhos sobre os modelos estatísticos clássicos de análise e previsão de séries temporais.

2.1 Pandemia de Coronavírus – COVID-19

A pandemia de COVID-19 se propagou de forma muito rápida pelo planeta e, em particular, no Brasil.

A evolução do número de casos no país é apresentada na figura 1, onde pode-se notar a forte ascensão que ocorreu

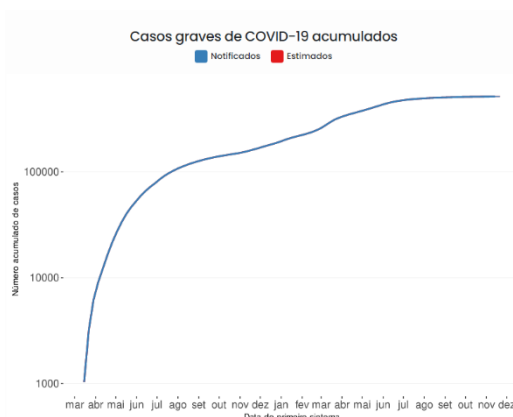


Figura 1 - Evolução do número de casos Brasil desde a primeira aparição em março de 2020
- escala logarítmica

Fonte: OBSERVATÓRIO COVID-19 BR

<https://covid19br.github.io/estados.html?uf=br>,

Há disponibilidade de grande quantidade de dados sobre a pandemia. A administração pública consegue levantar e tabular dados até o nível de distritos das cidades, e estas unidades geográficas, assim como, os municípios, são usados pelo Governo de São Paulo para elaborar suas diretrizes de isolamento e flexibilização da quarentena. O Laboratório Espaço Público e Direito à Cidade (LabCidade) da Faculdade de Arquitetura e Urbanismo da Universidade de São Paulo criou um mapa interativo que permite saber onde estão os casos graves confirmados e óbitos provocados pela covid-19 na Região Metropolitana de São Paulo. Esse grau de detalhamento possibilita análises mais precisas da difusão espacial da pandemia (EL PAÍS, 2020)

O enfrentamento está voltado para evitar uma rápida disseminação do vírus, reduzindo a ocorrência de contaminação de grande volume de pessoas de forma a não comprometer os sistemas de saúde. Veja-se, por exemplo, o caso da ocupação de leitos no município de São Paulo. O gráfico da figura 2 mostra o que ocorreria sem o distanciamento social. Pelo gráfico vê-se que sem as medidas de controle, como o distanciamento social, higiene pessoal e uso de máscaras, os leitos hospitalares da

cidade de São Paulo já teriam sido 100% ocupados pelos pacientes de COVID-19 entre os dias 1 e 5 de abril.

O objetivo dessas estratégias é a busca pelo achatamento da curva de contaminação populacional, de forma, a que se consiga retardar seu pico, e assim, diminuir a pressão sobre o sistema de saúde. Com isto, ganha-se tempo para melhor se preparar para o enfrentamento dos períodos mais graves (ALBUQUERQUE et al., 2020)

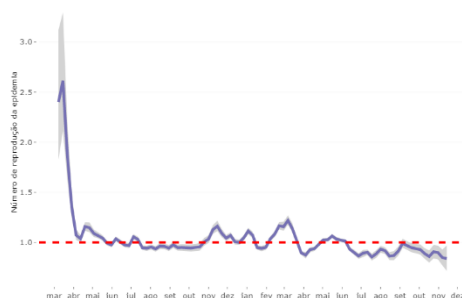


Figura 2 – Evolução do número de casos Brasil desde a primeira aparição em março de 2020
- escala logarítmica

Fonte: OBSERVATÓRIO COVID-19 BR

<https://covid19br.github.io/estados.html?uf=br>,

Para elaboração deste projeto, foram pesquisadas várias fontes de dados, em que se tem acesso às séries temporais do número de casos e de óbitos, que é o foco deste projeto. Dentre as fontes já citadas na seção 2.1, praticamente todas dão acesso a séries temporais de mais de uma variável associada à pandemia.

Além disso, a pesquisa mostrou que há ainda um volume considerável de trabalhos já desenvolvidos com acesso público, favorecendo o desenvolvimento deste projeto.

2.2 Análise e Previsão de Séries Temporais

Uma série temporal, também denominada série histórica, é uma sequência de dados obtidos em intervalos regulares de tempo durante um período específico. Na análise de uma série temporal, primeiramente deseja-se modelar o fenômeno estudado para, a partir daí, descrever o comportamento da série, fazer estimativas e, por último, avaliar quais os fatores que influenciaram o comportamento da série, buscando definir relações de causa e efeito entre duas ou mais séries. Para tanto, há um conjunto de técnicas estatísticas disponíveis que dependem do modelo definido (ou estimado para a série), bem como do tipo de série analisada e do objetivo do trabalho.” (CARDOSO e LATORRE, 2001)

A hipótese básica que norteia a análise de séries temporais é de que há um processo causal relativamente constante, relacionado com o tempo, que influencia os resultados. O objetivo da análise de séries temporais é identificar padrões não aleatórios na série para permitir previsões sobre o futuro, orientando a tomada de decisões.” (REIS, 2020)

Dentre os componentes de uma série temporal, destacam-se: a sazonalidade que é fenômeno que ocorre regularmente em períodos relativamente fixos de tempo (LATORRE e CARDOSO, 2001). Outro fator chave é a tendência. Antunes e Cardoso (2015), por exemplo, desenvolveram um estudo com a série temporal do coeficiente de mortalidade na cidade de São Paulo - SP, de 1900 a 1994, deixa claro que essa medida evoluiu ao longo do século reduzindo sua magnitude; o que se configura em uma tendência decrescente. No caso da série de COVID- 19 no Brasil, figura 1, o que se visualiza é uma tendência crescente. E há também numa série temporal, um fator aleatório, sobre o qual não se tem controle, e que é responsável pelos erros de previsão (MORETTIN e TOLOI, 1981; BOX e JENKINS, 1976).

A modelagem proposta por Box e Jenkins (1976) tornou-se um clássico da análise e previsão de séries temporais. A proposta é um método sistemático de identificação, ajuste, verificação e uso de modelos de séries temporais integradas, que envolvem médias móveis e componentes auto-regressivos. São os chamados modelos ARIMA - *autoregressive integrated moving average*.

Há muitos trabalhos em várias áreas com o uso desses modelos. Lodetti e Teixeira (2017), por exemplo, mostram uma aplicação no setor de vendas de leite, tendo relatado que obtiveram bons resultados.

Na área de epidemiologias, há aplicações também. Exemplos são os trabalhos de Antunes e Cardoso (2015) e de Latorre e Cardoso (2001). Em ambos os casos os autores relatam que obtiveram bons resultados.

3. METODOLOGIA

Esta foi uma pesquisa de natureza aplicada, com uma abordagem quantitativa, que utilizou como meios bibliografia, bases de dados, ferramentas computacionais e experimentos em ambiente de laboratório. Sua finalidade foi de desenvolvimento de uma metodologia para análise e previsão de séries temporárias

Após uma revisão bibliográfica do tema, para a análise dos dados do índice de isolamento das principais cidades do estado de São Paulo, foi utilizado como fonte os dados disponibilizados pelo governo do estado de São Paulo, através do site do SIMI-SP - Sistema de Monitoramento Inteligente de São Paulo (www.saopaulo.sp.gov.br/coronavirus/isolamento/), que é viabilizado por meio de acordo com as operadoras de telefonia Vivo, Claro, Oi e TIM, através da ABR (Associação Brasileira de Recursos em Telecomunicações) e do IPT (Instituto de Pesquisas Tecnológicas), para que o Estado possa consultar informações agregadas e anônimas sobre deslocamento nos municípios paulistas mapeados.' (ISOLAMENTO, 2021)

Segundo o SIMI, em respeito à proteção de dados, as informações são aglutinadas e anonimizadas, respeitando a privacidade dos usuários. Apresentando dessa forma, dados georreferenciados agrupados para elaborar políticas públicas que aprimorem as medidas de isolamento social para o enfrentamento ao coronavírus." (ISOLAMENTO, 2021).

Para essa análise foram utilizados os dados de índice de isolamento a partir do dia 26/02/2020 até o dia 23/04/2021.

Após a captura dos dados em uma planilha, foi necessário fazer alguns ajustes, a transformação dos dados que estavam em formato de texto para número, e a remoção dos acentos e da cedilha, para que fosse possível desenvolver qualquer análise com os dados, pela tecnologia R.

Além do levantamento dos dados, as seguintes etapas foram desenvolvidas no estudo:

- . Revisão Bibliográfica
- . Levantamento de Dados
- . Pré processamento de Dados
- . Aplicação de Algoritmos por meio de desenvolvimento de Scripts em R
- . Desenvolvimento de Experimentos
- . Análise de Resultados e Documentação

4. RESULTADO E DISCUSSÃO

Esta seção apresenta os experimentos desenvolvidos ao longo do estudo, as tecnologias utilizadas, os resultados obtidos e uma discussão sobre esses resultados

4.1 Pré processamento de Dados

Antes do desenvolvimento das análises foi necessário um processo de limpeza, transformação e integração dos dados na própria planilha Excel em que são fornecidos pelo SIMI-SP.

Foi preciso selecionar todos os dados que eram necessários formatar para número, mas que estavam em texto, e, após a seleção deles, bastava apenas digitar “DEC”.

Nos casos das cedilhas e acentos que tiveram que ser removidos no arquivo Excel, foi preciso criar uma função, que remove acentos. A criação de funções é feita pressionando-se ALT + F11, o que faz surgir uma nova janela do editor Visual Basic.

No menu acesse INSERIR/MÓDULO, foi construída a função descrita abaixo:

Function TiraAcento(Palavra)

```
CAcento = "àáâãäåèéêëìíîïòóôõöùúûüÿÀÁÂÃÄÅÈÉÊËÌÍÎÏÒÓÔÕÖÙÚÛÜÝÇçÑñ"
```

```
SAcento = "aaaaaeeeeiiiiooooouuuuuAAAAEEEEIIIOOOOOUUUUUcCnN"
```

```
Texto = ""
```

```
if Palavra <> "" then
```

```
For X = 1 to Len(Palavra)
```

```
Letra = mid(Palavra,X,1)
```

```
Pos_Acento = inStr(CAcento,Letra)
```

```
if Pos_Acento > 0 then
```

```
Letra = mid(SAcento,Pos_Acento,1)
```

```
end if
```

```
Texto = Texto & Letra
```

```
next
```

```
TiraAcento = Texto
```

```
end if
```

```
end function
```

```
Function VerificaPalavra(atributo)
```

```
Dim i
```

```
Dim id
```

```
Dim Auxiliar
```

```
Dim Resultado
```

```
Auxiliar = Split(Atributo, " ", - 1, vbBinaryCompare)
```

```
For i = LBound(Auxiliar) To Ubound(Auxiliar)
```

```
Resultado = Resultado & " " & TiraAcento(Auxiliar(i))
```

```
Next
```

```
VerificaPalavra = Trim(Resultado)
```

```
end function
```

E, para finalizar, foi preciso “girar” toda a base de dados. Na base original de Isolamento cada linha representa uma série temporal. Já na ferramenta R, que foi utilizada nas análises, os dados de uma série temporal são lidos como colunas. Portanto, foi preciso transformar as linhas da base de dados original em colunas, ou seja, foi preciso

construir uma matriz transposta dos dados originais. Isto foi feito no próprio R por meio de uma função.

4.2 Ferramentas Computacionais Utilizadas

A análise foi desenvolvida com o apoio de ferramentas computacionais disponíveis na linguagem R e em suas bibliotecas de algoritmos de séries temporais e de visualização de dados.

Em particular, foi usado o pacote “prophet”, que implementa um algoritmo que projeta uma série temporal fazendo uso de um modelo matemático do tipo aditivo, em que tendências não lineares são ajustadas aos dados e a estas são adicionadas sazonalidades, anual, semanal e diária. O pacote é considerado robusto para tratar dados ausentes, variações em tendências e *outliers* (CRAN.R, 2021a). Além disso, outra biblioteca importante foi aquela do pacote ‘forecast, que tem implementados algoritmos para análise e previsões de séries temporais, e visualização de resultados. Inclui o método de suavização exponencial e médias móveis e modelos clássicos de séries temporais (CRAN.R, 2021b). E uma última biblioteca que merece destaque é aquela do pacote ‘ggplot2, que possibilita a criação de gráficos, baseados na chamada gramática de gráficos (COX,2007), que trata da questão dos gráficos de forma ampla. No caso deste pacote, dá-se o acesso à base de dados, e informa-se como mapear as variáveis de interesse e definir a estética desejada. Com isto, muitas possibilidades de geração de gráficos estão disponíveis.

4.3 Desenvolvimento dos Experimentos

Conforme mencionado anteriormente, tinha-se em mãos series temporais do índice de isolamento das pessoas para os municípios do estado de São Paulo, durante um período de um ano entre fevereiro de 2020 e fevereiro de 2021. Para esta análise foram selecionadas as principais cidades das 16 regiões administrativas do estado de São Paulo, divididas em 4 sub-regiões: a capital, a região metropolitana de São Paulo (RMSP), o litoral e o interior do estado.

Em cada cidade escolhida foram gerados cinco tipos de análise:

- . Diagrama de dispersão dos dados originais do Índice de Isolamento;
- . Médias Móveis de 5, 7 e 10 dias;
- . Análise de Tendência;
- . Análise de Sazonalidade semanal e anual
- . Previsões com respectivo intervalo de confiança ao nível de 95%.

A média móvel é uma média dos valores de períodos que vão se sobrepondo. Assim, para uma média móvel de 5 dias, tem-se uma primeira média para os dias 1 a 5. A segunda média considera os dias 2 a 6. A terceira, considera os dias 3 a 7, e assim por diante.

A tendência é uma curva que identifica na curva dos dados originais tendências de crescimento, queda ou estabilização e representa esses comportamentos em diferentes períodos, por meio de uma curva.

Já a sazonalidade identifica movimentos de alta ou baixa para períodos específicos bem definidos no tempo. Uma sazonalidade semanal, por exemplo, identifica esses movimentos por dia da semana.

Para todas as cidades, tem-se uma primeira figura com um gráfico com 4 linhas: uma vermelha com os dados originais e nas linhas verde, azul e roxo, as médias móveis de 5, 7 e 10 dias respectivamente.

Em uma segunda figura, tem os gráficos de tendência, sazonalidade semanal e anual. Na sazonalidade semanal, apresenta o comportamento do índice de isolamento ao longo dos dias da semana, e no caso da sazonalidade anual mostra as variações mensais no índice de isolamento.

Uma terceira figura, apresenta os dados originais e as previsões, mostrando também as margens de erro, com nível de confiança estatístico de 95%. E, finalmente, uma última figura, apresenta um detalhamento das previsões, tendo-se três linhas: a linha vermelha com as previsões, linha azul com o limite superior de 95% das previsões e a linha verde que apresenta o limite inferior de 95%.

A seguir são apresentados esses resultados para sete cidades principais, representativas das regiões estudadas, acompanhados de comentários.

4.3.1 Análises de São Paulo (Capital)

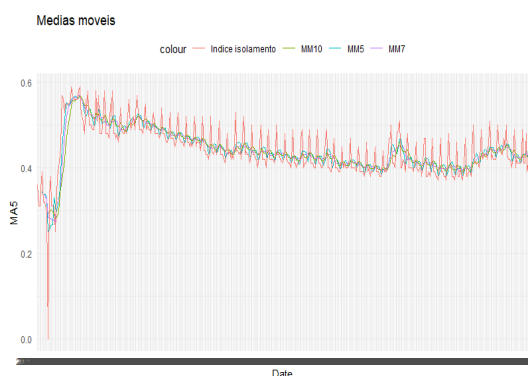


Figura 4.1a – Série Original e Médias Móveis
São Paulo – SP

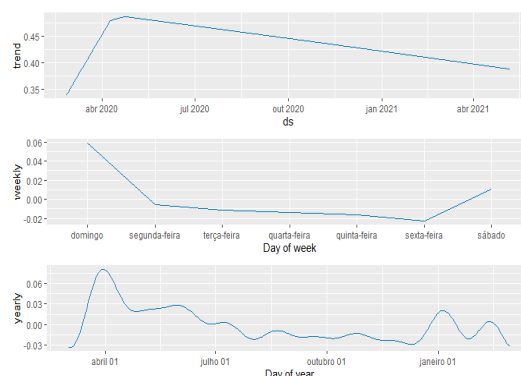


Figura 4.1b – Decomposição da Série
Tendência e Sazonalidade (semanal/anual)

São Paulo - SP

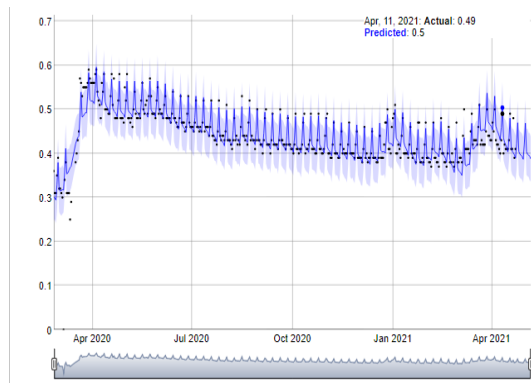


Figura 4.1c – Série Original e Previsões
São Paulo - SP

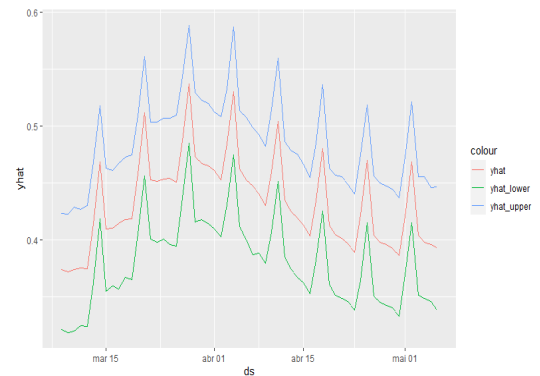


Figura 4.1d – Detalhe dos Intervalos de previsões
São Paulo - SP

Na figura 4.1a, nota-se que ao início da pandemia, muito por falta de informação de como tratá-la, o índice era baixo, porém ainda no início, as pessoas foram adquirindo informações de como se portar diante de uma pandemia, fazendo assim, com que o índice de isolamento aumentasse visto que uma das melhores alternativas para diminuição do contágio da COVID-19 era o distanciamento e isolamento social.

Na figura 4.1b, o gráfico de tendência demonstra no período, um aumento rápido do índice de isolamento entre o mês de fevereiro de 2020 até o mês de abril de 2020, e depois uma queda do final do mês de abril de 2020 até o mês de maio de 2021. Através do índice semanal pode-se concluir que durante o final de semana o índice de isolamento foi maior do que os índices dias de meio de semana, e, na sazonalidade anual, confirmar o que foi mostrado no gráfico 4.1a, que, já que em ambos os gráficos, o mês de janeiro de 2021, demonstrou um leve aumento do índice que vinha caindo mês após mês, e o mês de abril de 2020, foi o ápice do isolamento.

No gráfico 4.1c, diferente do que foi mostrado na tendência no gráfico 4.1b, teve um leve aumento, e durante o tempo de previsão, mesmo tendo uma variação, continuou tendo um índice consideravelmente alto.

E, por último, o gráfico 4.1d, assim como já mostrado no gráfico 4.1c, o índice tende a subir no início, porém mesmo com variações tende ter uma leve queda.

4.3.2 Análises da Região Metropolitana de São Paulo

- Guarulhos

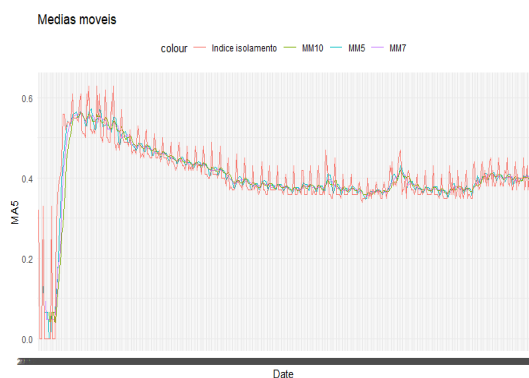


Figura 4.2a – Série Original e Médias Móveis
Guarulhos – SP

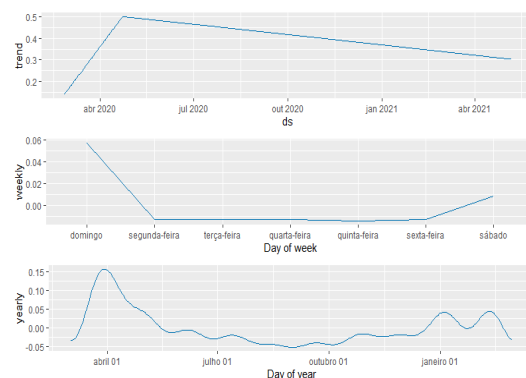


Figura 4.2b – Decomposição da Série
Tendência e Sazonalidade(semanal/anual)
Guarulhos - SP

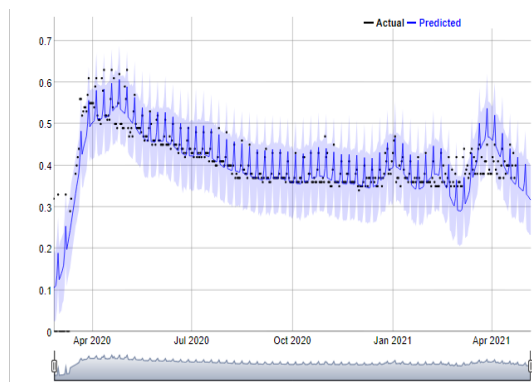


Figura 4.2c – Série Original e Previsões
Guarulhos – SP

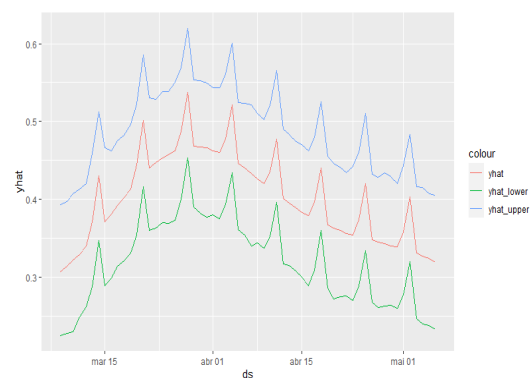


Figura 4.2d – Detalhe dos Intervalos de Previsões
Guarulhos - SP

Nas figuras 4.2a e na figura 4.2b, os gráficos tiveram o mesmo comportamento que na capital.

Já no gráfico 4.2c e no gráfico 4.2d, pode-se ver que diferente da tendência na figura 4.2b, no período da previsão teve-se um leve aumento no início da previsão, e antes da metade do período da previsão houve uma queda acentuada até o final dela.

- São Bernardo do Campo

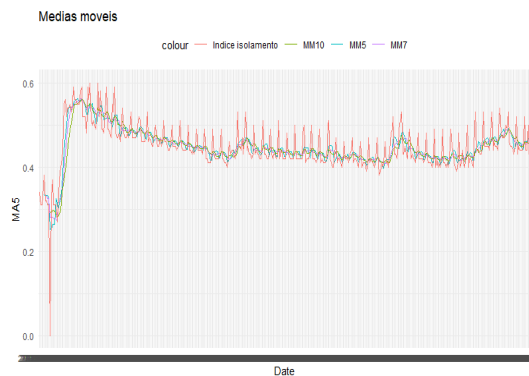


Figura 4.2e – Série Original e Médias Móveis
São Bernardo do Campo – SP

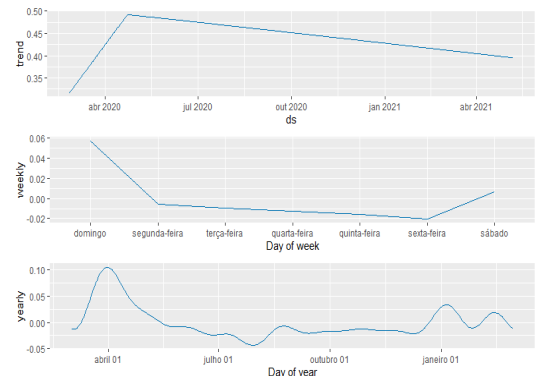


Figura 4.2f – Decomposição da Série
Tendência e Sazonalidade(semanal/anual)
São Bernardo do Campo - SP

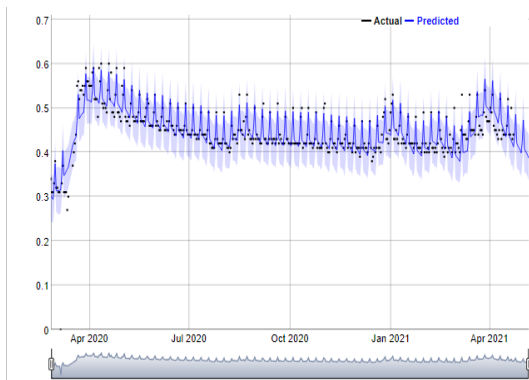


Figura 4.2g – Série Original e Previsões
São Bernardo do Campo – SP

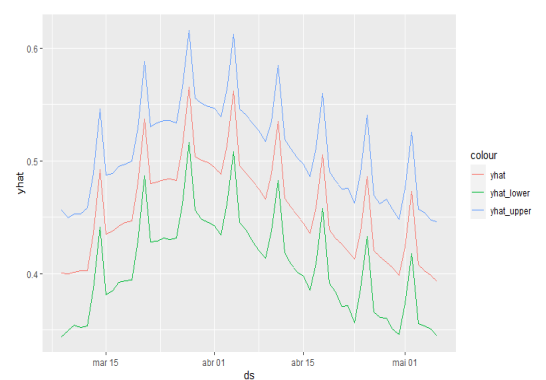


Figura 4.2h – Detalhe dos Intervalos de Previsões
São Bernardo do Campo - SP

Na cidade de São Bernardo do Campo, assim como nas cidades já citadas anteriormente, não se tinha tanta informação de como se comportar diante de uma pandemia, porém diferente dessas cidades já citadas ela teve um índice de isolamento maior e mais constante, como é demonstrado nas médias móveis da figura 4.3e.

Porém nas figuras 4.3f, 4.3g e 4.3h o comportamento dos gráficos demonstram as mesmas similaridades das cidades já citadas.

4.3.3 Análises do Litoral do Estado de São Paulo

- Santos

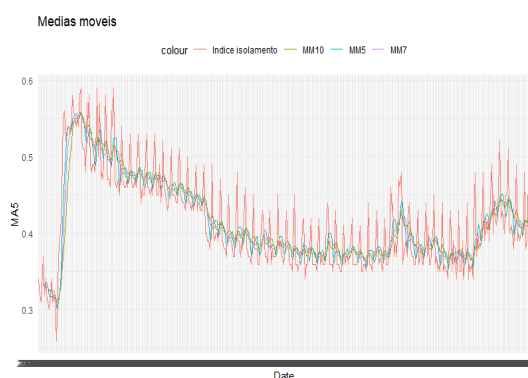


Figura 4.3a – Série Original e Médias Móveis Santos – SP

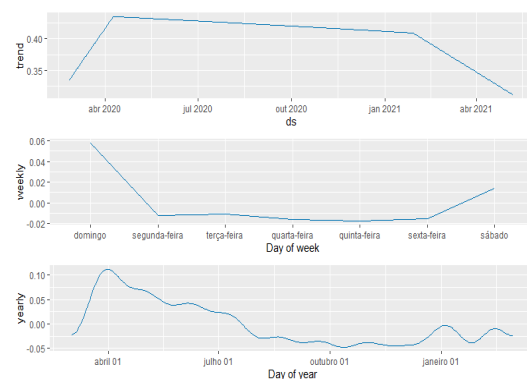


Figura 4.3b – Decomposição da Série Tendência e Sazonalidade(semanal/anual) Santos - SP

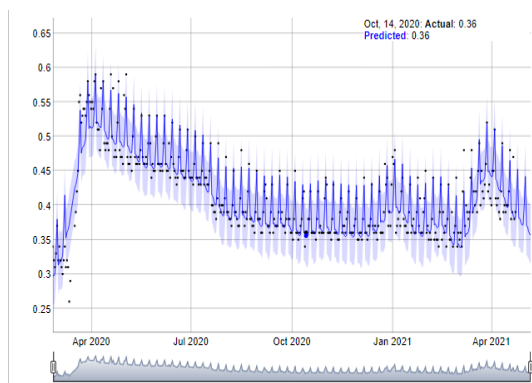


Figura 4.3c – Série Original e Previsões Santos – SP

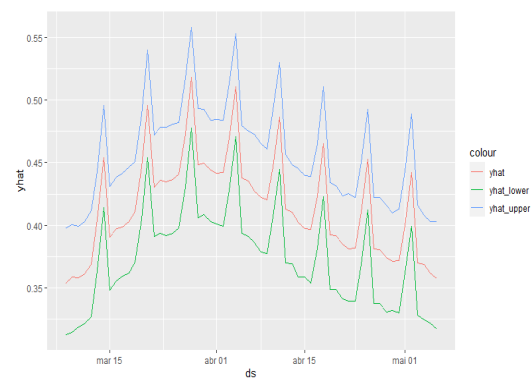


Figura 4.3d – Detalhe dos Intervalos de Previsões Santos - SP

Na cidade de Santos, notasse que o comportamento de isolamento demonstrado na figura 4.3a, teve o mesmo o início que nas outras cidades, porém a cidade de Santos teve uma queda mais brusca do índice de isolamento, do que nas cidades da região metropolitana e na capital.

Na figura 4.3b, o gráfico de tendencia demonstra, um aumento rápido do índice de isolamento entre o mês de fevereiro de 2020 até o mês de abril de 2020, e depois uma

queda do final de mês de abril de 2020 até o mês de fevereiro de 2021, porém após o mês de abril de 2021 ele tem uma queda brusca do índice de isolamento.

Na figura 4.3c, no período de previsão há uma certa peculiaridade, já que no início da previsão se tem um aumento considerável do índice de isolamento, e antes da metade do período da previsão esse índice despenca assim como nos dados originais do início da pandemia, como também pode-se observar na figura 4.3d.

4.3.4 Análises do Interior do Estado de São Paulo

- Campinas

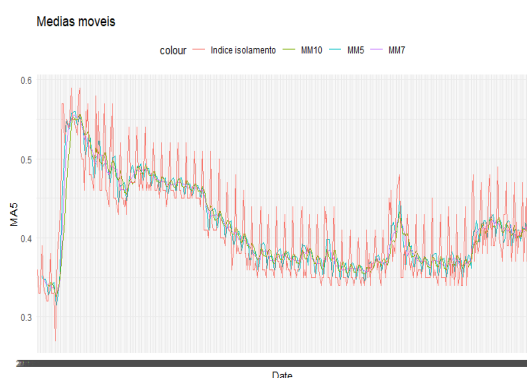


Figura 4.4a – Série Original e Médias Móveis
Campinas – SP

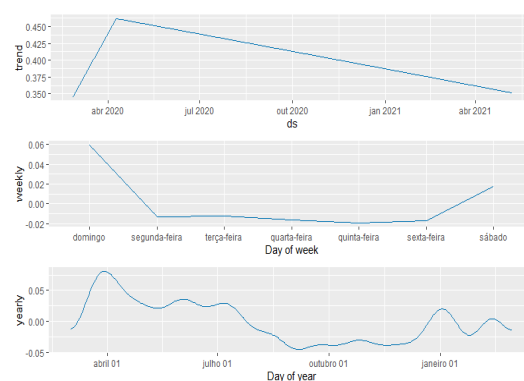


Figura 4.4b – Decomposição da Série
Tendência e Sazonalidade(semanal/anual)
Campinas - SP

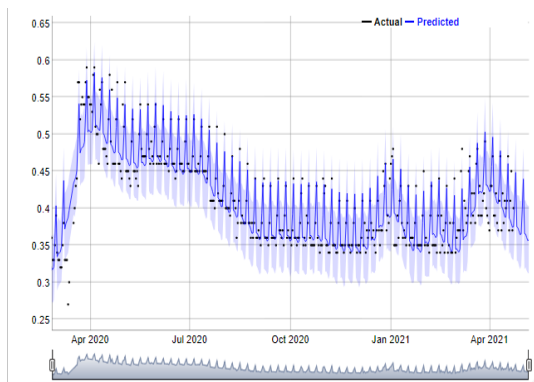


Figura 4.4c – Série Original e Previsões
Campinas – SP

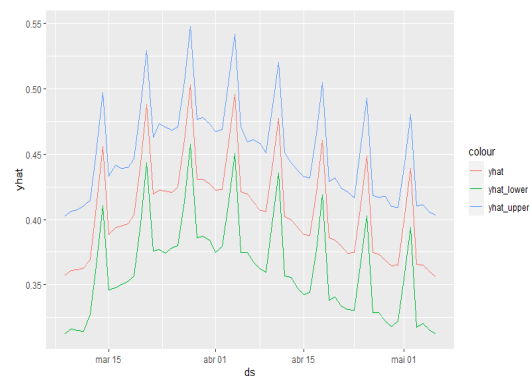


Figura 4.4d – Detalhe dos Intervalos de Previsões
Campinas - SP

Assim como na cidade de Santos, os dados originais da cidade de Campinas tiveram a mesma queda brusca do índice de isolamento, e um baixo valor do índice entre o período de Agosto de 2020 até Dezembro de 2020, que voltaram a cair no mês de Janeiro de 2021, como mostrado nas figuras 4.4a e 4.4c.

Na figura 4.4b, percebe-se que na tendência tem se um rápido aumento do índice de isolamento, como nas outras cidades, porém a partir do mês de abril de 2020 até abril de 2021 a tendência era de uma queda brusca do índice de isolamento.

E, em relação a previsão do índice de isolamento, a cidade de Campinas segue o mesmo padrão das outras cidades, tem uma leva crescente antes da metade do período, e depois uma queda do índice.

- Presidente Prudente

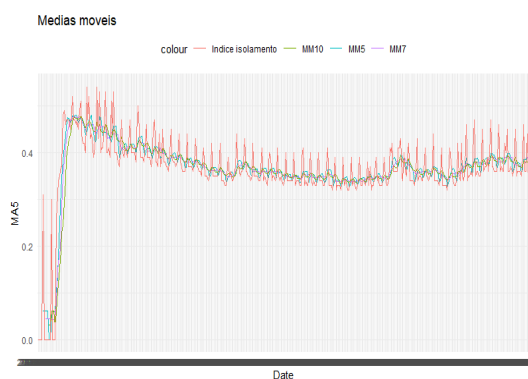


Figura 4.4e – Série Original e Médias Móveis Presidente Prudente – SP

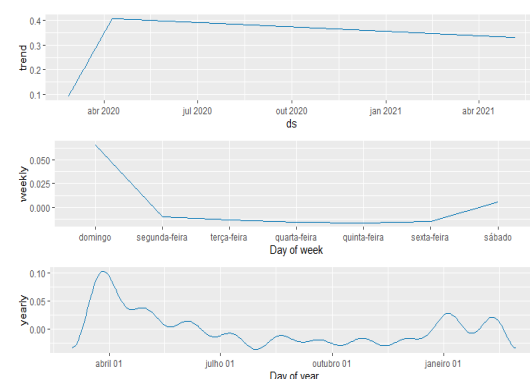


Figura 4.4f – Decomposição da Série Tendência e Sazonalidade(semanal/anual) Presidente Prudente - SP

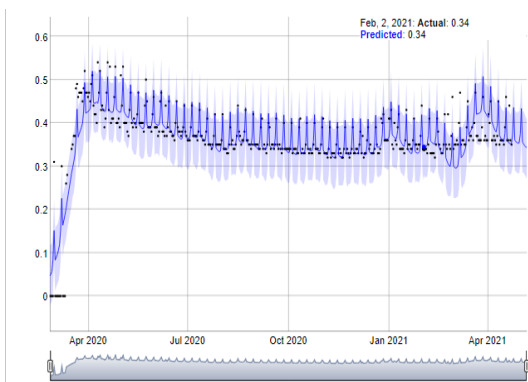


Figura 4.4g – Série Original e Previsões Presidente Prudente – SP

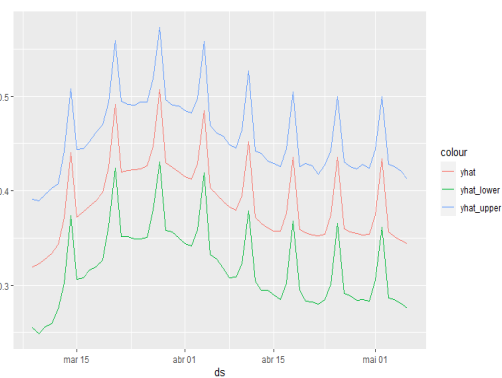


Figura 4.4h – Detalhe dos Intervalos de Previsões Presidente Prudente - SP

- Ribeirão Preto

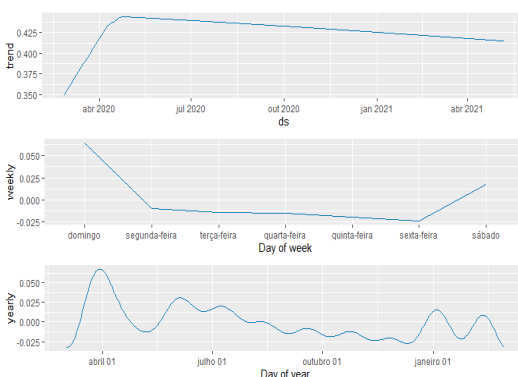


Figura 4.4i – Série Original e Médias Móveis
Ribeirão Preto – SP

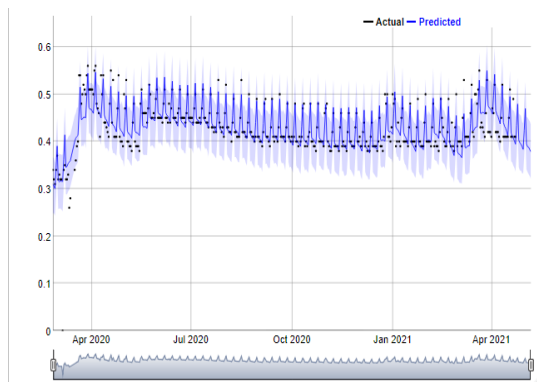


Figura 4.4k – Série Original e Previsões
Ribeirão Preto – SP

Figura 4.4j – Decomposição da Série
Tendência e Sazonalidade(semanal/anual)
Ribeirão Preto - SP

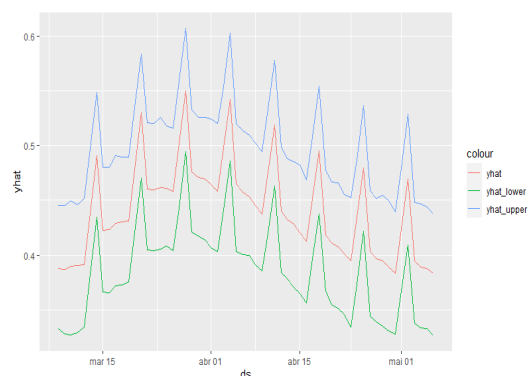


Figura 4.4l – Detalhe dos Intervalos de Previsões
Ribeirão Preto - SP

Diferentemente da cidade de Campinas, as cidades de Presidente Prudente e Ribeirão Preto, que representam o Interior, tiveram, a partir do mês de Abril, somente uma leve queda no índice, porém o índice ainda permanecia alto, como pode-se observar nas figuras 4.4e e 4.4i, o mesmo pode-se perceber na tendência do gráfico da figura 4.4f e 4.4j.

E nas previsões, como já dito, há o mesmo padrão, porém em comparação a cidade de Campinas o índice permanece alto.

5. CONSIDERAÇÕES FINAIS

Em resumo, conclui-se que através da sazonalidade semanal, todas as cidades tiveram nos dias de final de semana como os dias com maiores índices de isolamento. Além disso, nota-se que na série temporal dos dados originais mais as médias móveis, as cidades da região metropolitana de São Paulo, a capital e as cidades do interior, exceto a cidade de Campinas, tiveram no geral altos índices de isolamento, durante, ao menos um ano, em relação a cidade litorânea de Santos e a cidade de Campinas, que obtiveram baixos índice de isolamento, em comparação desse período.

Outro ponto a ser observado, foi que nas previsões geradas do índice de isolamento de todas as cidades apontadas no estudo, tiveram a mesma característica, um leve aumento do índice, e antes da metade do período ocorria uma queda.

REFERÊNCIAS

- ALBUQUERQUE, C. G. D.; COSTA, M. A.; CURI, R. L. C.; KRAUSE, C.; LUI, L.; SANTOS, R. M. D. S.; TAVARES, S. R. (2020). Apontamentos sobre a dimensão territorial da pandemia da COVID-19 e os fatores que contribuem para aumentar a vulnerabilidade socioespacial nas unidades e desenvolvimento humano de áreas metropolitanas brasileiras. Diretoria de Estudos e Políticas Regionais, Urbanas e Ambientais [Dirur], nº 15, p. 7.
- ALESSI, Gil e ROSSI, Marina (2020). Quantos de seus vizinhos em São Paulo contraíram o coronavírus? Mapa interativo da USP revela. El Pais, 11 de junho de 2020. Disponível em: https://brasil.elpais.com/brasil/2020-06-09/quantos-de-seus-vizinhos-em-sao-paulo-contrairam-o-coronavirus-mapa-interativo-da-usp-revela.html?ssm=FB_CC&fbclid=IwAR2wakC_oDhPGdGV0y8PJ8Nggwy5GMfrPJFHSA NBGc7w2sXNsRN6PE5PEbq
- BOX, G. E. P. e JENKINS, G. M. (1976) Time Series Analysis, Forecasting and Control. Holden-Day, San Francisco, Ca. 537 p.
- CRAN.R (2021a). Package prophet: Automatic Forecasting Procedure. The Comprehensive R Archive Network Project. Acesso em abril/2021. Disponível em : <https://cran.r-project.org/web/packages/prophet/index.html>
- CRAN.R (2021b). Package forecast: Forecasting Functions for Time Series and Linear Models. The Comprehensive R Archive Network Project. Acesso em abril/2021. Disponível em: <https://cran.r-project.org/web/packages/forecast/index.html>
- CRAN.R (2021c). Package ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics. The Comprehensive R Archive Network Project. Acesso em abril/2021, em: <https://cran.r-project.org/web/packages/ggplot2/index.html>

COX, N. J. (2007). Journal of Statistical Software. January 2007, Volume 17, Book Review 3. <http://www.jstatsoft.org/>

HYNDMAN, R.J., & ATHANASOPOULOS, G. (2018) Forecasting: principles and practice, 2nd edition, OTexts: Melbourne, Australia. Acessado em: janeiro/2021. Disponível em: <https://otexts.com/fpp2/>

Isolamento | Governo do Estado de São Paulo, Isolamento | Governo do Estado de São Paulo, disponível em: <<https://www.saopaulo.sp.gov.br/coronavirus/isolamento/>>, acesso em: 2 Aug. 2021

LATORRE, M. D. R. D. D. O.; CARDOSO, M. R. A. (2001). Análise de séries temporais em epidemiologia: uma introdução sobre os aspectos metodológicos. Time series analysis in epidemiology: an introduction to methodological aspects. Revista Brasileira de Epidemiologia. Vol. 4, Nº 3.

MINISTÉRIO DA SAÚDE (2020). (2020). Painel Coronavírus. <https://covid.saude.gov.br/>

MORETTIN, P. A.; TOLOI, C. M. C. (1981). Modelos para Previsão de Séries Temporais. IMPA – Instituto de Matemática Pura e Aplicada. Rio de Janeiro, RJ. 372 p.

OBSERVATÓRIO COVID-19 BR. COVID-19 no Brasil. <https://covid19br.github.io/>

REIS, M. M. (2000). Análise de séries temporais. Disponível em: <<https://www.inf.ufsc.br/~marcelo.menezes.reis/Cap4.pdf>> Acesso em: 10 de junho de 2020. Departamento de Informática e Estatística da Universidade Federal de Santa Catarina Campus Joinville – INE/UFSC.

Contatos: diego_2008cp@hotmail.com e aavallim@mackenzie.br