



**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE CIÊNCIAS DA SAÚDE
DEPARTAMENTO DE SAÚDE COLETIVA
CURSO DE GRADUAÇÃO EM SAÚDE COLETIVA**

YAN RODRIGUES ESTEVES

**Automação de alerta através do Diagrama de Controle na
identificação de surtos de dengue**

**BRASÍLIA-DF
2023**

YAN RODRIGUES ESTEVES

**Automação de alerta através do Diagrama de Controle na
identificação de surtos de dengue**

**Trabalho de Conclusão de Curso
como requisito parcial para
obtenção do título de Bacharel em
Saúde Coletiva**

Orientador: Ivan Zimmermann

Coorientador: Marcela Lopes Santos

BRASÍLIA-DF

2023

YAN RODRIGUES ESTEVES

**Automação de alerta através do Diagrama de Controle na
identificação de surtos de dengue**

**Trabalho de Conclusão de Curso
como requisito parcial para
obtenção do título de Bacharel em
Saúde Coletiva.**

BANCA EXAMINADORA

Henry Maia Peixoto

Felipe Ferré

Ivan Zimmermann

Aprovado em:

Brasília, _____ de _____ de _____.

Primeiramente, gostaria de enfatizar que sem essas pessoas eu não estaria concluindo este curso de Graduação e dedico a elas todo o sucesso que um dia terei por concluir essa importante jornada da minha vida.

Aos meus pais, Marcia e Junior. Irmãos, Maitê e Cauê. A melhor vó desse mundo, Helizete. E ao meu amor e futura esposa, Rebecca, que me apoiou e não me deixou desistir.

AGRADECIMENTOS

Aos professores que conheci nessa jornada, aos amigos que tive a oportunidade de desfrutar bons momentos em apresentações, trabalhos e pós-aulas, aos irmãos que a UnB me deu, a boa vontade dos Doutores que toparam participar da minha banca de defesa do TCC, a minha família e namorada, muito obrigado! Sem vocês eu não teria conseguido finalizar esta etapa da minha vida.

Marcela Lopes, muito obrigado por aceitar ser minha Coorientadora e por abrir as portas do MS, onde sempre foi meu sonho trabalhar. O convívio foi curto, mas o aprendizado ficará para sempre.

Alessandro do Qgiz, literalmente um irmão que se não fosse por conta da UnB, não existiram os Guerreiros do SUS. Obrigado pelas horas me ensinado sobre a estrutura de um trabalho acadêmico, horas de conversa jogada fora e momentos de se concentrar e manter a calma.

Ivan Zimmermann, obrigado pela confiança e incentivo em ampliar meu conhecimento nessa jornada da análise e visualização de dados através das inúmeras ferramentas que existem.

Felipe Ferré, comecei a me interessar por banco de dados e a entender um pouco mais desse universo graças a iniciativa e conversas que tivemos no CONASS, muito obrigado!

Maria Paula, muito obrigado por fazer parte da minha banca, e principalmente por lecionar IIS. Tomei gosto por esse mundo de gráficos e tabelas que aprendi graças a você.

Henry Peixoto, muito obrigado pela disponibilidade em ser meu suplente e fazer parte dessa fase da minha vida.

Rebecca Pissutti, se não fosse por você, nada disso teria acontecido. Obrigado por toda sua paciência, ajuda e compreensão. Não sei como será o futuro, mas quero você ao meu lado.

RESUMO

Considerando o alto número de casos de dengue no Brasil, este trabalho visa estruturar dados para a implementação de um alerta de surtos em processos de automação do Diagrama de Controle, utilizando a linguagem estatística R para auxílio da vigilância epidemiológica de dengue no Brasil. A partir de dados dos anos de 2015 a 2022 e com as bibliotecas em R *ggplot2*, *pacman*, *rio*, *tidyverse*, *dplyr*, *read.dbc*, *zoo* e *lubridate* e suas respectivas funções, foi possível moldar e calcular tabelas para gerar o alerta de cada Unidade Federativa do Brasil. Ao rodar o *script* no R, foram encontrados dois resultados, sendo o segundo resultado, o produto principal deste trabalho. Na primeira, obteve-se 27 gráficos de Diagramas de Controle, sendo um para cada estado e o DF, com dados da primeira semana epidemiológica (SE) até a SE 42. O segundo resultado gerado pelo *script* foi uma tabela, por SE de todos os estados e o DF, responsável por informar se o Coeficiente de Prevalência (CP) de 2022 foi ou não maior que a mediana dos anos anteriores e maior ou menor do que o Q3 por SE, através dos atributos “RISCO_mediana” e “RISCO_Q3”, mostrando de maneira mais simplificada quando o CI ultrapassa a mediana e o Q3, além de fornecer agilidade para a vigilância identificar possibilidade de surto e servindo como verificação do processo de criação dos gráficos. Após inspeção manual, pode-se observar que os resultados eram válidos. Nesse contexto, constata-se que o estudo atingiu seu objetivo principal, enquanto demonstrou como a automação pode auxiliar no controle de taxas de prevalência de dengue. Por conseguinte, a pesquisa contribuiu para a literatura dos principais temas nela abordados, isto é, a facilidade alcançada na vigilância epidemiológica por meio da automação de análises preditivas. A partir do código aqui implementado é possível automatizar os alertas, assim como replicar sua estrutura em outras áreas de interesse na vigilância.

Palavras-chave: Análise Preditiva; Curvas de Prevalência; Dengue; Diagrama de Controle; Vigilância Epidemiológica.

ABSTRACT

Considering the high number of dengue cases in Brazil, this work aims to structure data for the implementation of an outbreak alert in automation processes of the Control Diagram, using the R statistical language to assist the epidemiological surveillance of dengue in Brazil. From data from the years 2015 to 2022 and with the R libraries ggplot2, pacman, rio, tidyverse, dplyr, read.dbc, zoo, and lubridate and their respective functions, it was possible to shape and calculate tables to generate the alert for each Brazilian Federative Unit. When running the script in R, two results were found, with the second result being the main product of this work. In the first, 27 Control Diagram graphics were obtained, one for each state and the Federal District, with data from the first epidemiological week (SE) to EW 42. The second result generated by the script was a table, by EW for all states and the Federal District, responsible for informing whether the Incidence Coefficient (CI) of 2022 was or was not greater than the median of the previous years and greater or less than the Q3 by EW, through the columns "RISCO_mediana" and "RISCO_Q3", showing in a more simplified way when the CI exceeds the median and the Q3, in addition to providing agility for the surveillance to identify the possibility of an outbreak and serving as a verification of the graph creation process. After manual inspection, it was observed that the results were valid. In this context, it can be seen that the study achieved its main objective, while demonstrating how automation can assist in controlling incidence rates of dengue. Consequently, the research contributed to the literature of the main themes addressed in it, that is, the ease achieved in epidemiological surveillance through automation of predictive analyses. From the code implemented here, it is possible to automate alerts, as well as replicate its structure in other areas of interest in surveillance.

Keywords: Predictive Analyses; Incidence Curves; Dengue; Control Diagram; Epidemiological surveillance.

LISTA DE ABREVIATURAS E SIGLAS

DENV – O vírus da dengue

PNVS - Política Nacional de Vigilância em Saúde

SUS - Sistema Único de Saúde

SINAN - Sistema de Informação de Agravos de Notificação

SMS - Secretaria Municipal de Saúde

DATASUS - Departamento de Informática do Sistema Único de Saúde do Brasil

SE - Semana Epidemiológica

CP - Coeficiente de Prevalência

MS - Ministério da Saúde

VE - Vigilância Epidemiológica

SVS - Secretaria de Vigilância em Saúde

CEV - Campanha de Erradicação da Varíola

OMS - Organização Mundial de Saúde

OPAS - Organização Pan-Americana de Saúde

FSESP - Fundação Serviços de Saúde Pública

SNVE - Sistema Nacional de Vigilância Epidemiológica

Q1 - Primeiro Quartil

Q3 - Terceiro Quartil

CGARB - Coordenação Geral de Arboviroses

LISTA DE GRÁFICOS

Gráfico 1 - Diagrama de Controle de dengue, São Paulo, SE 42 de 2022.

Gráfico 2 - Diagrama de Controle de dengue, Acre, SE 42 de 2022.

Gráfico 3 - Diagrama de Controle de dengue, Alagoas, SE 42 de 2022.

Gráfico 4 - Diagrama de Controle de dengue, Amapá, SE 42 de 2022.

Gráfico 5 - Diagrama de Controle de dengue, Amazonas, SE 42 de 2022.

Gráfico 6 - Diagrama de Controle de dengue, Bahia, SE 42 de 2022.

Gráfico 7 - Diagrama de Controle de dengue, Ceará, SE 42 de 2022.

Gráfico 8 - Diagrama de Controle de dengue, Distrito-Federal, SE 42 de 2022.

Gráfico 9 - Diagrama de Controle de dengue, Espírito Santo, SE 42 de 2022.

Gráfico 10 - Diagrama de Controle de dengue, Goiás, SE 42 de 2022.

Gráfico 11 - Diagrama de Controle de dengue, Maranhão, SE 42 de 2022.

Gráfico 12 - Diagrama de Controle de dengue, Mato Grosso do Sul, SE 42 de 2022.

Gráfico 13 - Diagrama de Controle de dengue, Mato Grosso, SE 42 de 2022.

Gráfico 14 - Diagrama de Controle de dengue, Minas Gerais, SE 42 de 2022.

Gráfico 15 - Diagrama de Controle de dengue, Pará, SE 42 de 2022.

Gráfico 16 - Diagrama de Controle de dengue, Paraíba, SE 42 de 2022.

Gráfico 17 - Diagrama de Controle de dengue, Paraná, SE 42 de 2022.

Gráfico 18 - Diagrama de Controle de dengue, Pernambuco, SE 42 de 2022.

Gráfico 19 - Diagrama de Controle de dengue, Piauí, SE 42 de 2022.

Gráfico 20 - Diagrama de Controle de dengue, Rio de Janeiro, SE 42 de 2022.

Gráfico 21 - Diagrama de Controle de dengue, Rio Grande do Norte, SE 42 de 2022.

Gráfico 22 - Diagrama de Controle de dengue, Rio Grande do Sul, SE 42 de 2022.

Gráfico 23 - Diagrama de Controle de dengue, Rondônia, SE 42 de 2022.

Gráfico 24 - Diagrama de Controle de dengue, Roraima, SE 42 de 2022.

Gráfico 25 - Diagrama de Controle de dengue, Santa Catarina, SE 42 de 2022.

Gráfico 26 - Diagrama de Controle de dengue, Sergipe, SE 42 de 2022.

Gráfico 27 - Diagrama de Controle de dengue, Tocantins, SE 42 de 2022.

LISTA DE FIGURAS

Figura 1 - Quantidade de memória utilizada durante a leitura das tabelas e quantidade dos atributos.

Figura 2 - Atributos selecionadas de cada ano após a utilização da função *SELECT()*.

Figura 3 - Tabela indicativa de surto.

LISTA DE TABELAS

Tabela 1 - Tabela Verdadeiro ou Falso para CI maior que mediana e Q3 em São Paulo.

SUMÁRIO

1. INTRODUÇÃO	14
1.1 Objetivo Geral	16
1.2 Objetivos Específicos	16
2. REFERENCIAL TEÓRICO	17
2.1 Vigilância epidemiológica	17
2.2 Sistema de Informações de Agravos e Notificações (SINAN)	19
2.3 Análise preditiva	20
3. MÉTODOS	22
3.1 Contexto	22
3.2 Fontes de dados	22
3.3 Análise estatística	22
4. RESULTADOS E DISCUSSÕES	24
4.1 Código de execução em R	24
4.2 Diagramas de Controle	28
5. CONCLUSÃO	33
6. REFERÊNCIAS	34
APÊNDICE A. Diagramas de Controle de dengue nas unidades federativas do Brasil.	37

1. INTRODUÇÃO

Em 2009, a Organização Mundial de Saúde apresenta a dengue como um dos principais problemas de saúde pública no mundo, estando entre as doenças transmitidas pelo vetor, *Aedes aegypti*, que tem se espalhado geograficamente mais rápido nos últimos 50 anos. Com isso, cerca de 2,5 bilhões de pessoas vivem em áreas epidêmicas de dengue e ocorre anualmente cerca de 50 milhões de casos em todo o mundo (WORLD HEALTH ORGANIZATION, 2009). Com alta dispersão mundial, a Organização Pan-americana de Saúde, estima que 500 milhões de pessoas estão em risco de contrair a dengue e, em 2019, foram registrados mais de 3 milhões de casos, sendo 28 mil graves e 1.534 óbitos (PAN AMERICAN HEALTH ORGANIZATION, 2020).

O mosquito relacionado à transmissão da dengue é a fêmea *Aedes aegypti*, que, ao picar e consumir o sangue de uma pessoa infectada pelo vírus, pode transmitir para outros hospedeiros. A dengue possui quatro sorotipos, o DEN-1, DEN-2, DEN-3 e DEN-4 (OPAS, 2008) e pode se manifestar de diversas formas a depender do hospedeiro, podendo ser assintomática, ou até em casos graves e se não tratado, vir a óbito. Possui três fases clínicas, aguda, sistêmica e de recuperação (MINISTÉRIO DA SAÚDE, 2022a). Outras formas menos comuns de adquirir dengue, mas que ocorrem, são por meio transfusão de sangue, transplante de órgão e por relações sexuais desprotegidas (MINISTÉRIO DA SAÚDE, 2022a).

A variação de suas formas de apresentação pode incluir consequências graves, como a febre hemorrágica. Segundo a Organização Pan-americana de Saúde (2008), nas Américas, entre 2001 a 2007, foram notificados 106.037 casos de dengue hemorrágica, sendo o número total de mortes por dengue hemorrágica de 1.299, com uma taxa de mortalidade de 1,2% (PAN AMERICAN HEALTH ORGANIZATION, 2008). A dengue, em sua manifestação grave, é uma das causas de morte mais relevantes entre crianças em países da América Latina e da Ásia (PAN AMERICAN HEALTH ORGANIZATION, 2020).

O Brasil faz parte da sub-região conhecida como Cone Sul, em conjunto com outros países, como a Argentina, o Chile, o Paraguai e o Uruguai. No período de 2001 a 2007, 64,6% (2.798.601) de todos os casos de dengue nas Américas foram notificados nesta sub-região, dos quais 6.733 foram por dengue hemorrágica com um total de 500 óbitos (WORLD HEALTH ORGANIZATION, 2009). Cerca de 98,5%

dos casos foram notificados pelo Brasil, que também registra o maior número de casos e taxa de mortalidade na sub-região (WORLD HEALTH ORGANIZATION, 2009). Estudos mais recentes, realizados pelo Butantan em 2022, mostram que a região Centro-Oeste apresenta a maior taxa de incidência, com 920,4 casos de dengue por 100 mil habitantes, seguido pela região Sul que é a segunda maior com 472,2 casos de dengue por 100 mil habitantes, em terceiro a região Sudeste com 188,3 casos de dengue por 100 mil habitantes, o Norte com 154 casos de dengue por 100 mil habitantes e o Nordeste com 105 casos de dengue por 100 mil habitantes (INSTITUTO BUTANTAN, 2022a).

Desse modo, a vigilância epidemiológica de dengue no Brasil faz-se necessária para orientação técnica dos profissionais de saúde, para empregarem ações de controle, por meio de informações atualizadas sobre a sua ocorrência e agravos, bem como dos fatores que a condicionam, numa área geográfica ou população definida (MINISTÉRIO DA SAÚDE, 2005). Alguns termos-chave são relevantes neste contexto, incluindo pandemia, surtos, epidemias e endemias. A pandemia difere de uma situação de surto ou epidemia, pois atinge e se espalha em níveis globais, como ocorreu com a COVID-19. Um surto se caracteriza pelo aumento inesperado de casos em determinado período em um local restrito, enquanto uma epidemia é um aumento de casos em várias regiões, estados ou cidades. Por outro lado, uma endemia é uma situação recorrente em uma região, mas sem um aumento significativo de casos. (INSTITUTO BUTANTAN, 2022b) (BRASIL, 2018).

Paralelamente, variadas tecnologias em saúde são utilizadas para dar embasamento científico nas tomadas de decisão e possibilitar a execução de ações específicas para cada nível de surto encontrado. Assim, metodologias preditivas são úteis para detecção de surtos, pois permitem a utilização de recursos somente no momento adequado. Nesse contexto, a utilização do Diagrama de Controle surge para auxiliar a vigilância epidemiológica a se antecipar nas medidas de controle contra a dengue, em casos de possíveis surtos. Sendo importante para cada estado conseguir ter conhecimento de como está a realidade em seu território (BORTMAN, 1998).

A automação da análise de dados na saúde é uma necessidade e se torna cada vez mais importante e diferencial para o dia a dia do profissional de saúde dominar esse tipo de ferramenta a fim de lidar com os dados coletados. Um

exemplo disso é o estudo feito por Santos *et al*, (2019), no qual é demonstrado as etapas da utilização de uma metodologia de *machine learning* em análises preditivas com a finalidade de prever óbitos em uma população de idosos. Facilitando o direcionamento de políticas públicas. Neste contexto, torna-se oportuno um sistema de alerta automatizado de Diagrama de Controle, por meio de programação computadorizada, para surtos de dengue, auxiliando na redução do tempo de análise e investigação epidemiológica na detecção dos casos onde há a possibilidade de ocorrência de surto de dengue.

1.1 Objetivo Geral

O objetivo deste estudo é propor uma estrutura de dados utilizando a linguagem estatística R que permita auxiliar a análise de dados de notificação de casos de dengue registrados no Sistema de Informações de Agravos e Notificações (SINAN) de modo a auxiliar a vigilância epidemiológica de dengue no Brasil.

1.2 Objetivos Específicos

- A. Detectar a evolução de surtos de dengue a partir de Diagramas de Controle;
- B. Compreender a análise de Diagramas de Controle para dengue em todas as unidades federativas;
- C. Identificar os pacotes do R necessários e apropriados ao tratamento de dados e construção dos diagramas de controle;
- D. Propor uma rotina de programação em R que estruture os dados de notificação de modo a permitir a construção de alertas de possíveis surtos.

2. REFERENCIAL TEÓRICO

A seguir, serão apresentados e desenvolvidos os principais conceitos que englobam a temática do estudo. Assim, três tópicos centrais irão dar seguimento às descrições. O primeiro, apresenta um breve histórico da evolução da vigilância epidemiológica no Brasil. Em seguida, há o desenvolvimento de informações acerca do Sistema de Informações de Agravos e Notificações (SINAN). Por fim, os fundamentos teóricos sobre análise preditiva serão descritos.

2.1 Vigilância epidemiológica

A prática da vigilância epidemiológica iniciou antes do aparecimento de seu termo, portanto, no início do século XX, momento em que surgem os estudos sobre as bactérias e descobertas dos ciclos epidemiológicos de doenças parasitárias e infecciosas, começaram esforços para se intervir no ciclo dessas doenças a fim de diminuir e controlar a proliferação (BRASIL, 2005). Surge, então, o conceito de campanhas sanitárias, que tinham por objetivo combater vetores de doenças transmissíveis, evitando o adoecimento e também, não prejudicar a economia da época (BRASIL, 2005).

No Brasil, o Ministério da Saúde (2005) reconhece que a Campanha de Erradicação da Varíola (CEV), realizada entre 1966 e 1973, ficou reconhecida como o marco da institucionalização das ações de vigilância no país, fomentando a organização de unidades de vigilância epidemiológica na estrutura das secretarias estaduais de saúde, assim como auxiliando a Fundação Serviços de Saúde Pública (FSESP) a organizar, em 1969, um sistema de notificação semanal de doenças selecionadas e disseminar informações pertinentes em um boletim epidemiológico de circulação quinzenal.

Nesse contexto, é importante entender o conceito de Vigilância Epidemiológica adotado ao longo dos anos. Em sua percepção, Langmuir (1963, p. 182) define vigilância, quando aplicada a doenças, como sendo uma:

Observação contínua sobre a distribuição e tendências da incidência, mediante coleta sistemática, consolidação e avaliação de informes de morbidade e mortalidade, assim como de outros dados relevantes. Intrínseca ao conceito está a divulgação regular dos fundamentos, dados e interpretações para todos os que contribuíram e para todos os outros que precisam saber.

Com o passar dos tempos o conceito de vigilância foi se modelando e ganhando novos sufixos a fim de terem suas características diferenciadas e melhor definidas (ALBUQUERQUE et al., 2002). A adição do qualificativo “epidemiológica” ao termo “vigilância epidemiológica”, por exemplo, representou não apenas pesquisas direcionadas às doenças transmissíveis, mas a todos os tipos de doenças (ALBUQUERQUE et al., 2002).

Em 1975 ocorreu a 5ª Conferência Nacional de Saúde, responsável por colocar na Constituição Federal o Sistema Nacional de Vigilância Epidemiológica (SNVE), através da Lei nº 6.259/75, do Decreto nº 78.231/76 (BRASIL, 2005). Com isso, a notificação de doenças selecionadas na época tornaram-se obrigatórias e, desse modo, o Ministério da Saúde, mediante ao primeiro Guia de Vigilância de 1977, compartilha medidas técnicas para prevenção e controle de doenças transmissíveis emergentes no período que foram se adaptando para atender as demandas da época e evoluindo com o passar dos anos (BRASIL, 2005).

Na lei Nº 8.080 de 19 de setembro de 1990, o Sistema único de Saúde (SUS) coloca em prática a Vigilância epidemiológica através do SNVE, estabelecendo a normativa sobre o que se entende por vigilância epidemiológica:

Um conjunto de ações que proporcionam o conhecimento, a detecção ou prevenção de qualquer mudança nos fatores determinantes e condicionantes de saúde individual ou coletiva, com a finalidade de recomendar e adotar as medidas de prevenção e controle das doenças ou agravos.

Ainda, o Ministério da Saúde (2005) compreende que a vigilância epidemiológica tem por objetivo a orientação técnica para os profissionais de saúde, que, no que lhe concerne, se responsabilizam pelas ações de controle dessas doenças e agravos, por informações atualizadas sobre a ocorrência de doenças e agravos, bem como dos fatores que as condicionam, numa área geográfica ou população definida.

Complementarmente, existe um rol de atividades a serem realizadas pela vigilância, que se dividem em quatro tópicos, ao que compete a vigilância epidemiológica voltada para a dengue, segundo o Ministério da Saúde (2022b): (I) investigar as notificações de dengue e acompanhá-las periodicamente, com a intenção de identificar o comportamento das notificações durante um período; (II)

relacionar por análises epidemiológicas descritivas com padrões de tempo, espaço e população; (III) juntar os dados laboratoriais gerados pelas vigilâncias entomológicas; e (IV) diminuir a taxa de transmissão, trabalhando na prevenção das arboviroses a partir do compartilhamento de conhecimentos entre as esferas de vigilância.

2.2 Sistema de Informações de Agravos e Notificações (SINAN)

Para dar suporte na investigação e obtenção de dados para análises, em 1993, foi criado o Sistema de Informações de Agravos e Notificações (SINAN). Foi moldado com o passar dos anos, trazendo modelos de armazenamento de informações, apoiando no processo de investigação epidemiológico nas três esferas governamentais, que devem funcionar em rede (PORTAL SINAN, 2013).

A dengue é uma doença de notificação compulsória, ou seja, assim que detectada, independente de ser um caso suspeito, confirmado ou não confirmado, é obrigatório para o serviço de saúde notificá-la imediatamente. Assim que notificado, fica a cargo da Vigilância Epidemiológica (VE) tomar as ações necessárias quanto ao monitoramento dos casos e elaborar ações estratégicas para serem aplicadas à realidade de cada município (BRASIL, 2009).

Este processo ocorre utilizando os dados do SINAN nas fichas individuais de notificação e investigação que contém as informações básicas para as tabulações que serão realizadas pela VE municipal, que deve notificar e finalizar o acompanhamento de cada notificação de dengue, sendo obrigatório preencher o campo de encerramento da investigação em até 60 dias após a data de notificação (BRASIL, 2009).

Também fica a cargo da VE municipal investigar todos os casos notificados e montar um perfil dos bairros para repassá-los para a VE da Secretaria Estadual, responsável por fiscalizar os municípios e auxiliá-los nesse processo, acompanhando, identificando curvas epidêmicas, capacitando as equipes dos municípios, estabelecendo diretrizes e normas baseadas nos dados coletados para melhorar a vigilância, minimizando os danos da doença e repassar os dados ao Ministério da Saúde que deve fiscalizar no SINAN, se as informações estão sendo atualizadas constantemente, produzir boletins para disseminação da informação que

servirão para a criação de diretrizes nacionais para serem aplicadas com base na realidade de cada estado e município (BRASIL, 2009).

2.3 Análise preditiva

A análise preditiva consiste na utilização de algoritmos com a finalidade de compreender a estrutura de um banco de dados e como produto, obter regras de predição (REIS; BARCELLOS; PEDROSO; XAVIER, 2018). Essas regras são muito úteis na aplicação em indicadores de saúde, pois permitem, por exemplo, a possibilidade de prever eventos na saúde, como os riscos de determinado evento ocorrer, previsões de surtos e endemias. Isso dependerá do tipo e quantidade de dados econômicos, sociais, demográficos, e outros tipos de dados que se possui (SANTOS *et al.*, 2019). Sendo muito útil, também, para o embasamento quantitativo na construção de políticas públicas, permitindo aos gestores não apenas saber como está a realidade do território, mas sendo possível realizar análises e planejar conforme as necessidades, a fim de melhorar a realidade e qualidade de vida daquele território (SANTOS *et al.*, 2022).

Outro estudo que defende a utilização de métodos preditivos de surtos é o de Pafiadache, *et al.* (2015), que demonstra por gráficos de controle, inicialmente utilizados no setor da indústria e adaptados à saúde. Sendo possível utilizar desses mesmos métodos, adaptados para o cenário da saúde, a fim de prever um evento. No estudo em questão, foram utilizados os dados do SINAN e dois métodos de detecção estatísticos, os Gráficos X e MR (para medidas individuais e amplitude móvel) e medidas exponencialmente ponderáveis (EWMA). A objetividade do estudo era saber qual dos métodos desempenha melhor para uma detecção de surto, sendo o EWMA demonstrando ser mais eficiente. Ressaltando a importância para a vigilância epidemiológica de se ter uma base de dados confiável, dado a importância de se prever corretamente os eventos (PAFIADACHE, *et al.*, 2015).

Desse modo, a Análise Preditiva torna-se adequada para o cumprimento dos objetivos do atual estudo, permitindo realizar a predição do quadro da dengue nos estados do Brasil. Uma das maneiras de realizar esta predição, é por meio do Diagrama de Controle que utiliza de probabilidade estatística para visualizar o comportamento dos anos anteriores a partir dos dados de um evento utilizando o

limite máximo, mínimo e a mediana, sendo assim, quando os resultados dessas medidas estão distantes do limite superior e da mediana, apresentam uma probabilidade menor de ocorrerem e quando os resultados tocam a mediana ou ultrapassam, a probabilidade de ocorrerem aumentam (ARANTES *et al.*, 2003).

3. MÉTODOS

3.1 Contexto

Para este estudo serão consideradas as notificações de dengue nas unidades federativas brasileiras no intervalo de 2015 a 2022. Foi utilizado também a população estimada do IBGE para os anos de 2015 a 2022, disponíveis no DATASUS.

3.2 Fontes de dados

Os dados referentes à notificação de dengue de 2015 a 2021 foram adquiridos no DATASUS e os de 2022, foi disponibilizado pela Coordenação de Arbovirose do Ministério da Saúde (CGARB), pois ainda não estavam disponíveis no DATASUS. Será utilizado também a população estimada do IBGE para os anos de 2009 a 2022, disponíveis no DATASUS.

3.3 Análise estatística

O presente trabalho possui um caráter descritivo. Com a finalidade de alcançar o objetivo deste trabalho, foram reproduzidas análises preditivas baseadas no Diagrama de Controle para a detecção de surtos de dengue no Brasil, utilizando a linguagem de análises estatísticas R. O ambiente R traz uma grande variedade de ferramentas, pacotes e técnicas estatísticas, permitindo tratar os dados a partir de sua linguagem de programação (ABOUT R, 2022), que pode ser replicada e, caso necessário, automatizada.

Por se tratar de um software livre, não é necessário pagar por uma licença para utilização, a comunidade ativa é expressiva e está em constante crescimento, permitindo auxílio em fóruns online, como o *stack overflow*, e sua replicação se limita apenas ao tempo necessário para aprender a linguagem, recursos e uso dos pacotes necessários para cada finalidade (STACKOVERFLOW, 2022). Foram aplicados os pacotes *pacman*, *rio*, *tydeverse*, *dplyr*, *zoo* e *lubridate*. Estes pacotes são necessários para o tratamento e limpeza das tabelas, a fim de padronizar, organizar, torná-las estruturadas para melhor funcionamento do software R, promovendo credibilidade às informações. Esses dados tratados formam uma única

tabela, que, na última etapa, será utilizada como fonte para gerar os gráficos, através do pacote *ggplot*.

A automação neste processo se dá uma vez que o *script* esteja construído, tornando a reprodução mais simples e otimizando o processo de vigilância epidemiológica. O detalhamento do *script* utilizado, o qual é um texto com instruções para o software R do que se deseja realizar com os dados que se possui, será descrito a seguir.

4. RESULTADOS E DISCUSSÕES

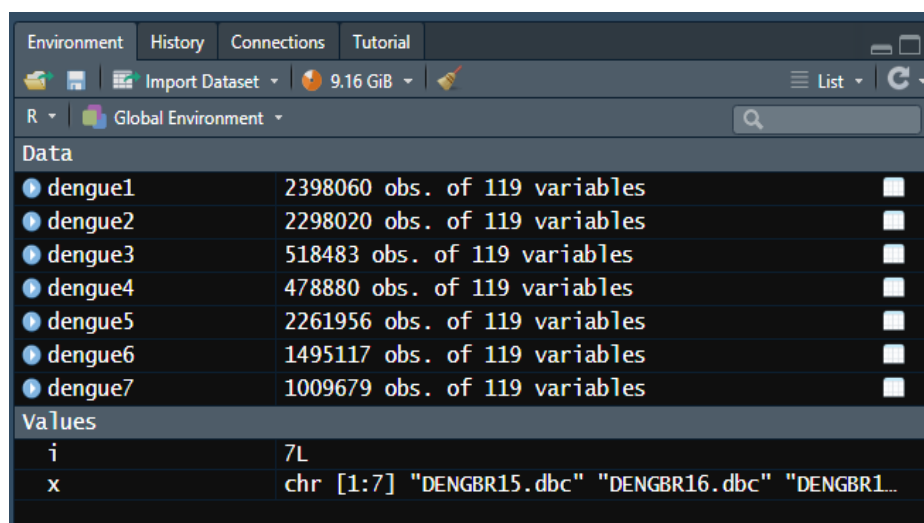
4.1 Código de execução em R

Após a condução de testes preliminares, obteve-se a proposta de um roteiro de análise em R:

O primeiro passo para começar a análise é ter todos os dados que serão utilizados pelo *software* estatístico R, como as tabelas de dengue de 2015 a 2022, assim como a população extraída do IBGE nesses respectivos anos.

O segundo passo é baixar e instalar as bibliotecas de pacotes que serão utilizadas para a leitura, tratamento e junção dos dados. Na leitura foram utilizados os pacotes, *read.dbc* responsável por permitir a leitura de arquivos que estejam em formato DBC (comum em arquivos do DATASUS) e o pacote *readxl* para a leitura de arquivos em formato Excel (.xlsx) (Batra *et al.*, 2021). Em seguida, foi armazenado em uma variável todo o dado de 2015 a 2021 de dengue. Esse processo exige muita memória RAM do computador, sendo possível visualizar na figura 1, pois todos os dados coletados da ficha do SINAN foram armazenados para se ter acesso rápido a eles.

Figura 1: Quantidade de memória utilizada durante a leitura das tabelas e quantidade dos atributos.



The screenshot shows the RStudio Environment pane with the following data objects and their memory usage:

Object	Memory Usage
dengue1	2398060 obs. of 119 variables
dengue2	2298020 obs. of 119 variables
dengue3	518483 obs. of 119 variables
dengue4	478880 obs. of 119 variables
dengue5	2261956 obs. of 119 variables
dengue6	1495117 obs. of 119 variables
dengue7	1009679 obs. of 119 variables

Below the data objects, the 'Values' section shows the following information:

Variable	Value
i	7L
x	chr [1:7] "DENGBR15.dbc" "DENGBR16.dbc" "DENGBR1..."

Fonte: Elaborado pelo autor

O terceiro passo é excluir os atributos que não foram necessárias para o cálculo. Quando as tabelas foram extraídas possuíam 119 atributos, ver gráfico 1, representando todos os campos de observação da ficha do SINAN para dengue. Após a aplicação do pacote *dplyr*, responsável pela transformação e manipulação de dados, utiliza-se a função *select*, que seleciona os atributos desejados, as quais são DT_SIN_PRI (data dos primeiros sintomas), CLASSI_FIN (informa a classificação final do caso) e ID_MN_RESI (código IBGE do município de residência), ver figura 2. Para cada ano é realizada essa alteração e ao final é utilizado o comando básico do R, o *rbind*, responsável por juntar as linhas que, neste caso, uniu todas as informações de 2015 a 2021 em uma variável.

Figura 2: Atributos selecionadas de cada ano após a utilização da função SELECT()

Variable Name	Number of Observations	Number of Variables
dengue1	2398060	3
dengue2	2298020	3
dengue3	518483	3
dengue4	478880	3
dengue5	2261956	3
dengue6	1495117	3
dengue7	1009679	3

Variable	Value
i	7L
x	chr [1:7] "DENGBR15.dbc" "DENGBR16.dbc" "DENGBR1..."

Fonte: Elaborado pelo autor

O quarto passo foi unir os dados de 2022 aos dados de 2015 a 2021. No decorrer deste trabalho, o ano de 2022 estava em curso, logo, os dados para este ano não estão completos. O processo é o mesmo do terceiro passo, seleciona-se apenas três atributos e ao final une-se o ano de 2022 com os anos de 2015 a 2021, com a utilização da função *rbind*. Ao final nomeia-se esses dados unificados de “dengue”.

No quinto passo aplica-se novamente o pacote *readxl* para a leitura da população total estimada pelo IBGE de 2009 a 2022. Para este passo, utiliza-se o

termo pivotagem, pois a tabela vem no formato longo, quando há mais linhas do que atributos, e através da função *pivot_longer()*, da biblioteca *tidyr*, ela foi transformada para o formato largo, onde há mais atributos do que linhas (CURSO-R, 2023). E ainda dentro dessa função, foram selecionados os atributos que serão utilizados, respectivamente aos anos de 2015 a 2022, pela função *cols =*. Como foi alterado o formato da tabela, é necessário dizer quais os novos nomes dos atributos, pela função *names_to =*, e da primeira linha *values_to =*.

O sexto passo é responsável por verificar a variável das datas, para saber como está seu formato na tabela, se realmente está no formato de data e por qual estrutura de data se encontra, para poder padronizar para o formato *aaaa-mm-dd*.

O sétimo passo se verifica as categorias da variável dengue para excluir os casos descartados, que na tabela do SINAN essa coluna pode vir codificada, sendo o objetivo encontrar as linhas que vêm com o código "5", pois essas representam os casos descartados. Utiliza-se um filtro para excluir os registros onde possuir o "5" na coluna *class*.

O oitavo passo é o momento de agrupar todos os registros no dia em que ocorreram. Para isso, na tabela "dengue", criada ao final do 4º passo, utiliza-se somente a coluna "uf", representando o código da Unidade Federativa. Utilizando a biblioteca *dplyr* e suas funções, *group_by*, responsável por agrupar todos os dados que possuem o mesmo valor e a função *summarise*, responsável por criar uma tabela que será organizada com os diferentes agrupamentos de dados criados pela função *group_by* (Batra *et al.*, 2021).

O nono passo é o momento de mudar as categorias que estão como código numérico para o significado em palavras, que neste caso, é o nome de cada estado do Brasil. Para isso é necessário transformar a coluna "uf" em fator e denominar para cada numeração o seu respectivo nome. E ao final excluir as linhas onde não houverem registros, para padronizar e não ter erros de leitura.

O décimo passo é para organização das datas em semana epidemiológica, para isso, utilizaram-se os pacotes *lubridate* e *zoo*, combinados. O primeiro responsável por permitir o manuseamento e transformação da variável data (Batra *et al.*, 2021). Neste processo, agregando os dias em semanas. E o outro, é utilizado para trabalhar com períodos de data atípicos (Batra *et al.*, 2021). Como, por exemplo, semanas epidemiológicas, organizadas pelo primeiro dia sendo domingo e o último dia, sábado (BRASIL, 2021).

O décimo primeiro passo foi criar uma coluna para receber a população de cada ano, colocado anteriormente no quinto passo. Com a população de 2015 a 2022, somando com a tabela dengue. Podendo assim, seguir para o próximo passo.

O décimo segundo passo é o momento de calcular a prevalência de dengue, através da fórmula de prevalência, casos de dengue naquele período dividido pela população multiplicado por cem mil. Antes de prosseguir para o próximo passo, a forma de detectar o surto utilizando o Diagrama de Controle, neste trabalho é através da mediana calculada a partir da prevalência de casos de dengue dos anos anteriores (2015 a 2021) e comparadas com a curva de prevalência do ano atual (2022), por SE.

Sendo assim, o décimo terceiro passo foi o cálculo da mediana e dos quartis (Q1 e Q3). Para este cálculo foi utilizado um conjunto de funções na ordem: *filter()*, para realizar um filtro e excluir o ano 2022, pois neste momento, utilizaram-se somente os anos 2015 a 2021. *Mutate()*, para criar uma coluna com as datas por SE. *Group_by()* para executar os cálculos por grupo de SE e por Unidade Federativa de residência. *Summarise()*, para realizar os cálculos de mediana do CI, calcular o Primeiro Quartil (Q1) e o Terceiro Quartil (Q3). E utilizou-se *mutate()* por último para criar dois novos atributos com até duas casas decimais para o Q1 e o Q3. Este cálculo foi armazenado em uma variável chamada 'dengue12_21' e será utilizada no passo número 15.

O décimo quarto passo, é a utilização do pacote *ggplot2* para a criação dos gráficos, um para cada estado e Distrito Federal.

O décimo quinto passo é o momento de realizar uma verificação simplificada, onde é gerada uma tabela dizendo quais os estados que tiveram o coeficiente de prevalência maior ou menor que a mediana e o Q3 dos últimos anos. Para isso foram utilizadas as bibliotecas *dplyr* e *lubridate*, carregadas no primeiro passo. Utilizou-se duas tabelas já criadas, a primeira (dengue12_21) contém os cálculos da mediana e Q3. A outra (dengue2022_) possui o coeficiente de prevalência de dengue. Ambas possuem dois atributos em comum, a semana epidemiológica e a UF de residência, que serão utilizadas para unir esses dois atributos em um pela função *string_c()* (Batra *et al.*, 2021). Permitindo assim, a união das tabelas através da função *merge()*, que por semelhanças encontradas nas duas tabelas, uni todos os atributos das duas tabelas, formando uma, com todos os dados (Batra *et al.*, 2021).

4.2 Diagramas de Controle

Ao rodar o *script*, o produto final foi dividido em duas partes. Na primeira, obteve-se 27 gráficos de Diagramas de Controle, sendo um para cada estado e o DF, com dados da primeira semana epidemiológica (SE) até a SE 42. A análise foi feita a partir da curva da incidência gerada pelos dados do coeficiente de prevalência (CI), analisados semanalmente em 2022. É utilizado, igualmente, a mediana do CI e o Terceiro Quartil (Q3) dos anos de 2015 a 2021, também por SE, formando assim o Diagrama de Controle (BRASIL, 2022b). É importante reforçar que este processo já vem sendo realizado pela equipe de arboviroses do Ministério da Saúde, replicado neste estudo para proposição do produto no segundo resultado.

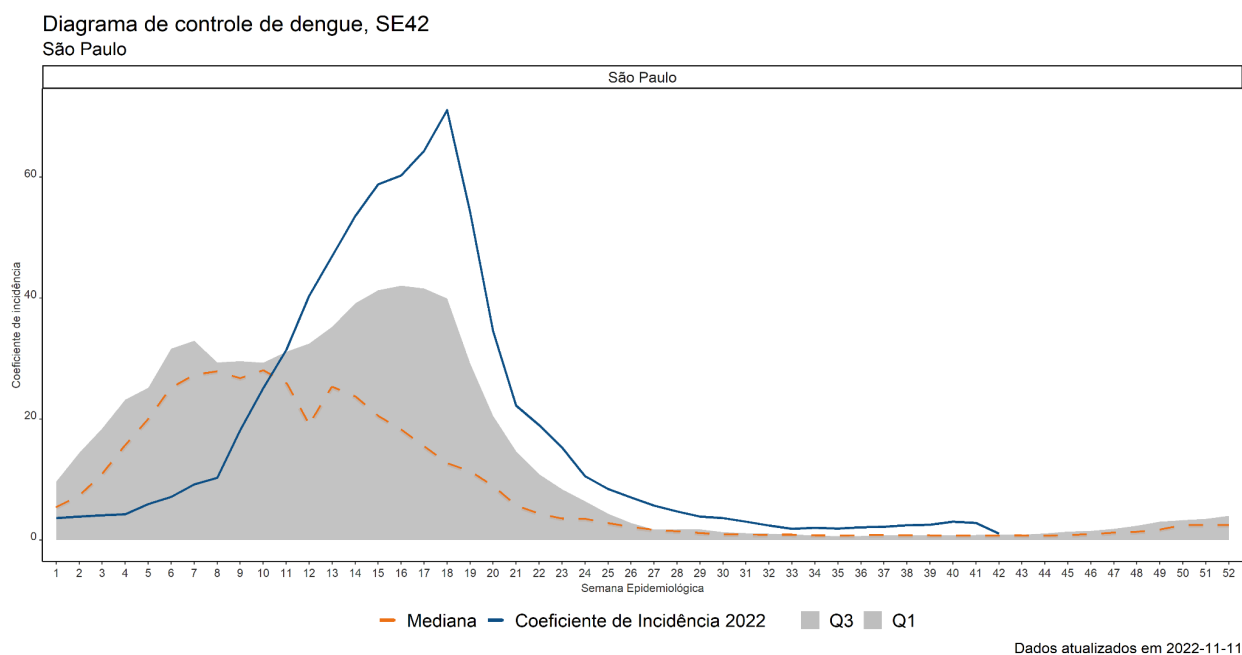
Para analisar os Diagramas de Controle, basta observar o CI de 2022, representado pela linha contínua azul. Se o CI ultrapassar a mediana dos últimos anos (2015 a 2021), representado pela linha intermitente laranja, entra-se em estado de alerta e deve ser notificado ao estado para serem tomadas as medidas adequadas (ARANTES *et al.*, 2003). Caso ultrapasse o Q3, o estado entra em situação de risco e as ações de vigilância devem ser intensificadas e aplicado o plano de contingência (BRASIL, 2022).

O Gráfico 1 representa o Diagrama de Controle do estado de São Paulo e será utilizado como exemplo para a explicação dos gráficos dos demais estados e DF (vide Apêndice A). Neste gráfico, é possível evidenciar que na primeira SE o CI estava menor, quando comparado à mediana dos anos anteriores. Nas SE 11 a 18 foi constatado um aumento no CI, ultrapassando a mediana. Ainda que este aumento fosse esperado, considerando o histórico da mediana e devido caracterizar-se como um período sazonal, época em que há aumento esperado de casos (BRASIL, 2022), há indicação da necessidade de alerta ao estado.

Assim, além de passar a mediana, também ultrapassou o Q3, representando, na SE 11, o início do surto de dengue no estado de São Paulo, atingindo seu pico na SE 18. Na SE 19 o CI começa reduzir, mas a todo momento, até a SE 42, se mantém acima da mediana e do Q3, permanecendo o cenário de alerta ao estado.

Neste trabalho, os dados vão até a SE 42 devido à execução do mesmo ter começado já no final de 2022, quando esses dados ainda não existiam.

Gráfico 1: Diagrama de Controle de dengue, São Paulo, SE 42 de 2022.



Fonte: Elaborado pelo autor

O segundo resultado gerado pelo *script* foi uma tabela, por SE de todos os estados e o DF, responsável por informar se o CI de 2022 foi ou não maior que a mediana dos anos anteriores e maior ou menor do que o Q3 por SE, através dos atributos “RISCO mediana” e “RISCO Q3”, resultando em uma análise mais precisa e servindo como verificação do processo de criação dos gráficos. Quando há o estado de alerta, observa-se “Verdadeiro” na coluna “RISCO mediana” e “Falso” na coluna “RISCO Q3”. Ao se observar “Verdadeiro” em ambos atributos “RISCO mediana” e “RISCO Q3”, representa-se a necessidade de intensificar a intervenção. Já o resultado “Falso” em ambos os atributos “RISCO mediana” e “RISCO Q3”, mostra não haver possibilidade de surto.

Como um complemento para vigilância de dengue, a Tabela 1 mostra, de maneira mais simplificada, quando o CI ultrapassa a mediana e o Q3, fornecendo agilidade para a vigilância identificar possibilidade de surto. Nesse exemplo, demonstra-se um recorte do estado de São Paulo das SE 8 a 27, onde se visualiza, especificamente, os momentos em que o CI ultrapassa a mediana e o Q3. É possível observar que esses momentos ocorrem a partir da SE 11, e se mantém até o final dos registros.

Tabela 1 - Tabela Verdadeiro ou Falso para CI maior que mediana e Q3 em São Paulo.

Semana Epidemiológica	UF Residência	Mediana	Q3	CI	RISCO mediana	RISCO Q3
8	São Paulo	27,77	29,39	10,32	FALSO	FALSO
9	São Paulo	26,64	29,57	18,17	FALSO	FALSO
10	São Paulo	27,95	29,33	25,12	FALSO	FALSO
11	São Paulo	25,79	31,14	31,43	VERDADEIRO	VERDADEIRO
12	São Paulo	18,87	32,495	40,37	VERDADEIRO	VERDADEIRO
13	São Paulo	25,21	35,27	46,98	VERDADEIRO	VERDADEIRO
14	São Paulo	23,58	39,16	53,61	VERDADEIRO	VERDADEIRO
15	São Paulo	20,4	41,325	58,81	VERDADEIRO	VERDADEIRO
16	São Paulo	18,09	42,075	60,3	VERDADEIRO	VERDADEIRO
17	São Paulo	15,36	41,625	64,33	VERDADEIRO	VERDADEIRO
18	São Paulo	12,57	39,95	71,11	VERDADEIRO	VERDADEIRO
19	São Paulo	11,22	29,22	54,19	VERDADEIRO	VERDADEIRO
20	São Paulo	8,8	20,54	34,59	VERDADEIRO	VERDADEIRO
21	São Paulo	5,57	14,68	22,23	VERDADEIRO	VERDADEIRO
22	São Paulo	4,21	10,89	18,98	VERDADEIRO	VERDADEIRO
23	São Paulo	3,44	8,405	15,34	VERDADEIRO	VERDADEIRO
24	São Paulo	3,39	6,42	10,56	VERDADEIRO	VERDADEIRO
25	São Paulo	2,67	4,35	8,43	VERDADEIRO	VERDADEIRO
26	São Paulo	2,03	2,84	7,05	VERDADEIRO	VERDADEIRO
27	São Paulo	1,59	1,75	5,72	VERDADEIRO	VERDADEIRO

Fonte: Elaborado pelo autor

Por fim, a Figura 3 é outra forma de visualizar o momento em que o CI ultrapassa a mediana e o Q3, também através dos atributos “RISCO mediana” e “RISCO Q3”, indicando um alerta por estado e a necessidade de intensificar as medidas de contingência. Quando há o estado de alerta, observa-se “True” na coluna “RISCO mediana” e “False” na coluna “RISCO Q3”. Ao se observar “True” em ambas os atributos “RISCO mediana” e “RISCO Q3”, representa-se necessidade de

intensificar a intervenção. Já o resultado “False” em ambos os atributos “RISCO mediana” e “RISCO Q3”, mostra não haver possibilidade de surto.

Neste caso, é analisada apenas a primeira SE de todos os estados e DF. Como a análise no Ministério da Saúde é realizada ao nível estadual por SE, a tabela é gerada dessa forma. Sendo assim, é possível analisar diretamente quando um estado não apresenta possibilidade de surto (15 estados), entra em alerta de possível surto (2 estados) ou necessita intensificar a intervenção para diminuição de casos (10 estados). Conforme relatado, esta análise é referente somente à primeira SE, sendo que o método de análise se repete para as demais semanas.

Figura 3 - Tabela indicativa de surto.

se.x	ufres.x	mediana	Q3	ci	RISCO_mediana	RISCO_Q3
1	Acre	45.160	67.0100	17.64	FALSE	FALSE
2	Amapá	1.170	3.3500	0.92	FALSE	FALSE
3	Amazonas	2.360	3.5050	1.52	FALSE	FALSE
4	Bahia	2.610	2.7050	2.38	FALSE	FALSE
5	Espírito Santo	9.190	12.8700	0.02	FALSE	FALSE
6	Maranhão	0.680	0.9600	0.54	FALSE	FALSE
7	Mato Grosso do Sul	7.360	21.1100	2.31	FALSE	FALSE
8	Minas Gerais	4.090	6.5100	2.73	FALSE	FALSE
9	Paraná	1.800	8.2650	0.99	FALSE	FALSE
10	Pernambuco	1.280	3.2200	0.77	FALSE	FALSE
11	Rio de Janeiro	1.720	2.3200	0.15	FALSE	FALSE
12	Rio Grande do Norte	3.360	5.8800	1.45	FALSE	FALSE
13	Roraima	0.950	2.6125	0.18	FALSE	FALSE
14	São Paulo	5.340	9.6750	3.63	FALSE	FALSE
15	Sergipe	0.260	1.9300	0.13	FALSE	FALSE
16	Alagoas	0.970	2.9200	1.86	TRUE	FALSE
17	Ceará	0.930	3.1000	1.28	TRUE	FALSE
18	Distrito Federal	5.090	7.3200	21.73	TRUE	TRUE
19	Goiás	17.710	29.4000	62.37	TRUE	TRUE
20	Mato Grosso	9.170	14.4250	30.62	TRUE	TRUE
21	Pará	0.940	1.6400	2.89	TRUE	TRUE
22	Paraná	1.430	2.1400	3.07	TRUE	TRUE
23	Piauí	1.020	1.4150	2.65	TRUE	TRUE
24	Rio Grande do Sul	0.085	0.1375	0.25	TRUE	TRUE
25	Rondônia	3.010	5.4950	13.23	TRUE	TRUE
26	Santa Catarina	0.160	0.2700	0.28	TRUE	TRUE
27	Tocantins	5.280	14.6750	76.57	TRUE	TRUE

Fonte: Elaborado pelo autor

Na análise preditiva não há apenas uma metodologia que funcione, existem várias, como as utilizadas por BORTMAN (1998), PAFIADACHE (2015), ARANTES (2003), SANTOS (2019) e SANTOS (2022). Apesar de cada uma possuir seu método, todas utilizam dos acontecimentos no passado para se preparar para o futuro.

Autores como BORTMAN (1998) comprovam que os métodos preditivos em análise de dados de saúde são úteis em mostrar como está a realidade do território, dando embasamento quantitativo para identificação de surtos. No estudo, ressalta a importância de reconhecimento do canal endêmico realizado pela média de casos mensais dos últimos 5 ou 7 anos, para identificar períodos de sazonalidade da doença (BORTMAN, 1998), que neste trabalho é realizado através do Diagrama de Controle. Sendo possível identificar essa previsão no Gráfico 1 em São Paulo nas SE 2 a 22, através do histórico gerado pelo Diagrama de Controle.

ARANTES (2003), utiliza um método também baseado na média de casos dos anos anteriores para prever surtos hospitalares em um estudo de coorte, chamado de distribuição probabilística de Poisson. Utilizado para amostras menores, possibilitando evidenciar padrões e prever quando um surto ocorre. (ARANTES *et al.*, 2003).

Ainda que o modelo e o delineamento de pesquisa tenham sido construídos cuidadosamente, foi possível identificar algumas limitações. Pode-se perceber que o trabalho foi construído em cima dos dados sobre a dengue, assim a generalização de suas análises para demais arboviroses se faz limitada, apesar de que o modelo de *script* realizado pode ser aplicado em análises de outras enfermidades. Assim, recomenda-se que estudos futuros desenvolvam a possibilidade de automação de análises preditivas de prevalência de demais arboviroses existentes, como na Zika, Chikungunya e Febre Amarela, por exemplo, para auxiliar no controle e no desenvolvimento de ações preventivas de surtos. Além disso, sugere-se que, para fins de melhor visualização dos dados tratados, seja realizada a transformação dos dados em dashboard interativo ou painel.

5. CONCLUSÃO

A utilização de métodos estatísticos de extração, transformação e visualização de dados são muito úteis na vigilância epidemiológica da dengue. Quando combinados com recursos computadorizados potencializam as ações da vigilância epidemiológica, permitindo sua replicação em qualquer nível, podendo ser modificado e melhorado para resolver cada situação.

O Diagrama de Controle é um dos métodos que, através de cálculos estatísticos baseados na mediana do Coeficiente de Prevalência dos últimos anos (2015 a 2021), permite observar o histórico da prevalência de casos dengue, e em conjunto com a curva de prevalência e do Coeficiente de Prevalência do ano atual (2022), prevê o início de surto, podendo auxiliar a tomada de decisão e direcionamento dos recursos para combater a dengue.

A tabela criada como segundo produto do *script* facilita ainda mais o processo, uma vez que, gera um alerta através dos atributos RISCO_mediana e RISCO_Q3, otimizando o processo para o profissional da vigilância, permitindo, assim, a detecção precoce dos surtos, possibilitando a ação da vigilância em tempo oportuno e diminuição de danos.

6. REFERÊNCIAS

ALBUQUERQUE, Maria Ilk Nunes de; CARVALHO, Eduardo M. Freese de; LIMA, Luci Praciano. **Vigilância epidemiológica: conceitos e institucionalização**. Revista Brasileira de Saúde Materno Infantil, v. 2, n. 1, p. 7–14, 2002. Acessado em: 22 dez. 2022.

ARANTES, Aglai et al. **Uso de diagramas de controle na vigilância epidemiológica das infecções hospitalares**. Revista de Saúde Pública [online]. 2003, v. 37, n. 6, pp. 768-774. Disponível em: <https://doi.org/10.1590/S0034-89102003000600012>. Epub 25 Mar 2004. ISSN 1518-8787. Acessado em: 6 dez. 2022.

BATRA, Neale, et al. **The Epidemiologist R Handbook**. 2021. Disponível em: <https://epirhandbook.com/en/#citation>. Acessado em: 01 jan. 2023.

BORTMAN, Marcelo. **Elaboración de corredores o canales endémicos mediante planillas de cálculo**. Rev Panam Salud Publica/Pan Am J Public Health 5(1), 1999. p.8. Disponível em: <<https://scielosp.org/article/rpsp/1999.v5n1/1-8/es/>> Acessado em: 06 dez. 2022.

BRASIL. Dive Santa Catarina. **Calendário epidemiológico**. 2021. Disponível em: <https://dive.sc.gov.br/phocadownload/calendario-epidemiologico/Calend%C3%A1rio%20Epidemiologico/2022.pdf>. Acessado em: 22 dez. 2022.

BRASIL. Instituto Butantan. **Veja como se prevenir da dengue; número de casos em 2022 já é maior que o total de 2021**. 2022a. Disponível em: <https://butantan.gov.br/noticias/veja-como-se-prevenir-da-dengue--numero-de-casos-em-2022-ja-e-maior-que-o-total-de-2021>. Acessado em: 07 dez. 2022.

BRASIL. Instituto Butantan. **Entenda o que é uma pandemia e as diferenças entre surto, epidemia e endemia**, Brasil. 2022b. Disponível em: <<https://butantan.gov.br/covid/butantan-tira-duvida/tira-duvida-noticias/entenda-o-qu-e-e-uma-pandemia-e-as-diferencas-entre-surto-epidemia-e-endemia>> Acessado em: 07 dez. 2022.

BRASIL. Lei n. 8.080, de 19 de set. de 1990. **Dispõe sobre as condições para a promoção, proteção e recuperação da saúde, a organização e o funcionamento dos serviços correspondentes e dá outras providências.**, Brasília, DF, set 1990. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l8080.htm. Acessado em: 10 nov. 2022.

BRASIL. Ministério da Saúde. **Diretrizes Nacionais para a Prevenção e Controle de Epidemias de Dengue**. 2009. Disponível em: https://bvsmms.saude.gov.br/bvs/publicacoes/diretrizes_nacionais_prevencao_controle_dengue.pdf. Acesso em: 04 dez. 2022.

BRASIL. Ministério da Saúde. **Guia de Vigilância em Saúde**, 2022a. Disponível em: https://bvsmms.saude.gov.br/bvs/publicacoes/guia_vigilancia_saude_5ed_rev_atual.pdf. Acessado em: 09 nov. 2022.

BRASIL. Ministério da Saúde. **Guia de Vigilância Epidemiológica**, 2005. Disponível em: https://bvsmms.saude.gov.br/bvs/publicacoes/Guia_Vig_Epid_novo2.pdf. Acessado em: 09 nov. 2022.

BRASIL. Ministério da Saúde. **Guia para Investigações de Surto ou Epidemias / Ministério da Saúde**, 2018. Disponível em: https://bvsmms.saude.gov.br/bvs/publicacoes/guia_investigacao_surto_epidemias.pdf. Acessado em: 09 nov. 2022.

BRASIL, Ministério da Saúde. **Plano de Contingência Para Respostas às emergências em Saúde Pública por Dengue, Chikungunya e Zika**. 2022b. Disponível em: <https://www.gov.br/saude/pt-br/centrais-de-conteudo/publicacoes/publicacoes-svs/dengue/plano-de-contingencia-para-resposta-as-emergencias-em-saude-publica-por-dengue-chikungunya-e-zika>. Acessado em: 11 jan. 2023.

CURSO-R. **O pacote lubridate**. Disponível em: <https://livro.curso-r.com/7-5-o-pacote-lubridate.html>. Acesso em: 01 jan. 2023.

CURSO-R. **O pacote tidyverse**. Disponível em: <https://livro.curso-r.com/4-2-tidyverse.html>. Acesso em: 01 jan. 2023.

LANGMUIR, Alexander D. The Surveillance of Communicable Diseases of National Importance. **New England Journal of Medicine**, v. 268, n. 4, p. 182–192, 1963. Acesso em: 11 nov. 2022.

PAFIADACHE, Caroline, ZANINI, Ruviano, SOUZA, Adriano Mendonça, DA ROSA, Leandro Cantorski, PERIPOLLI, Angélica. **Uma Aplicação de Gráficos de Controle Estatístico e Modelos de Previsão De Box-Jenkins para Monitoramento de Sistemas de Saúde no Rio Grande do Sul, Brasil**. Espacios, 36 (nº17) pág. 6. 2015. Acessado em: 22 dez. 2022.

PAHO, Organização Pan-Americana de Saúde, 2008. **Number of Reported Cases of Dengue & Dengue Hemorrhagic Fever (DHF), Region of the Americas (by country and subregion)**. Disponível em: <https://www3.paho.org/English/AD/DPC/CD/dengue-cases-2008.htm>. Acessado em: 09 nov. 2022.

PAHO, Organização Pan-Americana de Saúde, 2020. **Dengue**. Disponível em : <https://www.paho.org/pt/topicos/dengue>. Acessado em 07 dez. 2022.

PERGUNTAS frequentes. **Portal Sinan**, 2013. Disponível em: <http://www.portalsinan.saude.gov.br/perguntas-frequentes>. Acesso em: 12 nov. 2022.

R. **about R**. 2022. Disponível em: <https://www.r-project.org/about.html>. Acesso em: 30 nov. 2022.

REIS, Gilson Jácome dos; BARCELLOS, Christovam; PEDROSO, Marcel de Moraes; XAVIER, Diego Ricardo. **Diferenciais intraurbanos da sífilis congênita: análise preditiva por bairros do Município do Rio de Janeiro, Brasil**. 2018. Cadernos de Saúde Pública, Rio de Janeiro, v. 34, n. 9, 6 set. 2018. FapUNIFESP (SciELO). <http://dx.doi.org/10.1590/0102-311x00105517>. Acessado em: 15 dez. 2022.

SANTOS, Hellen Geremias dos; NASCIMENTO, Carla Ferreira do; IZBICKI, Rafael; DUARTE, Yeda Aparecida de Oliveira; FILHO, Alexandre Dias Porto Chiavegatto. **Machine learning para análises preditivas em saúde: exemplo de aplicação para prever óbito em idosos de São Paulo, Brasil**. Cadernos de Saúde Pública, 35, nº.7, 2019. Acessado em: 20 dez. 2022.

SANTOS, Natalia; SILVA, Everton; BELUZO, Carlos; ALVES, Luciana C.. **Modelos de Previsão para Avaliação de Casos de Malária na Amazônia Legal Brasileira**. In: ENCONTRO NACIONAL DE INTELIGÊNCIA ARTIFICIAL E COMPUTACIONAL (ENIAC), 19. , 2022, Campinas/SP. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2022 . p. 670-681. ISSN 2763-9061. DOI: <https://doi.org/10.5753/eniac.2022.227370>. Acessado em: 22 dez. 2022.

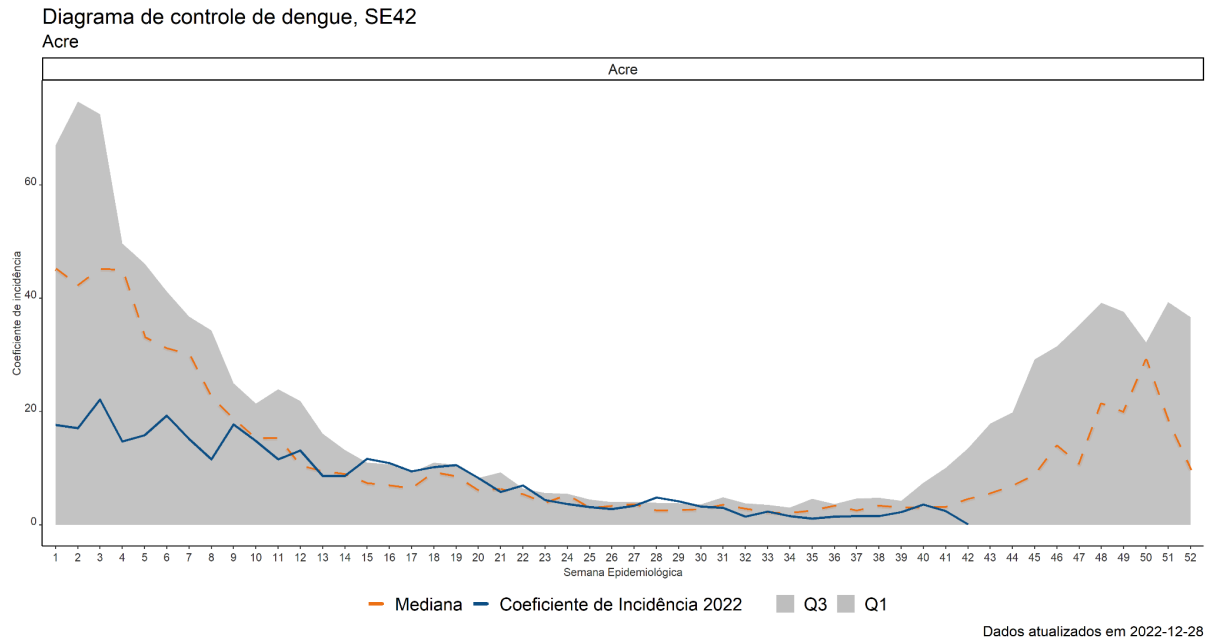
STACK OVERFLOW. **Who we are**, 2022. Disponível em: <https://stackoverflow.co/>. Acesso em: 30 nov. 2022.

WHO, Organização Mundial de Saúde, 2009. **DENGUE GUIDELINES FOR DIAGNOSIS, TREATMENT, PREVENTION AND CONTROL**. Disponível em: https://reliefweb.int/report/world/dengue-guidelines-diagnosis-treatment-prevention-and-control?gclid=CjwKCAiAyfybBhBKEiwAgtB7fqsD1kpkXgh8prV-kBuTAKbo9M8b3Ha0ek0J6fFdinRAIbtWRSAdxBoCYKkQAvD_BwE>. Acessado em: 24 nov. 2022.

APÊNDICE A.

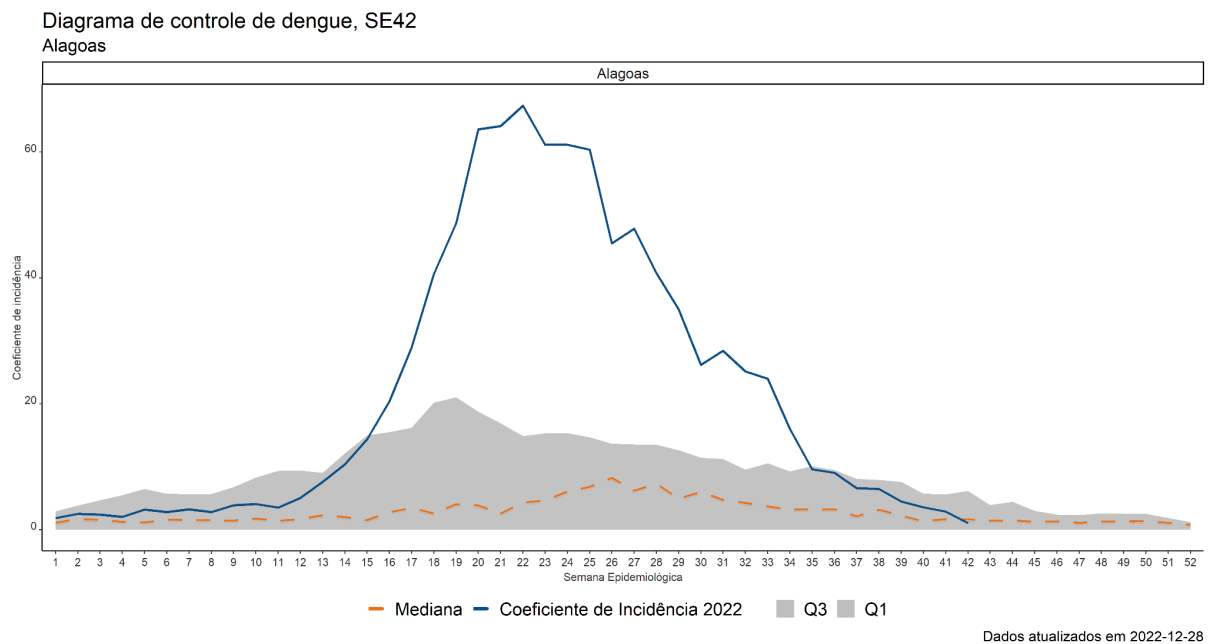
Diagramas de Controle de dengue nas unidades federativas do Brasil.

Gráfico 2 - Diagrama de Controle de dengue, Acre, SE 42 de 2022.



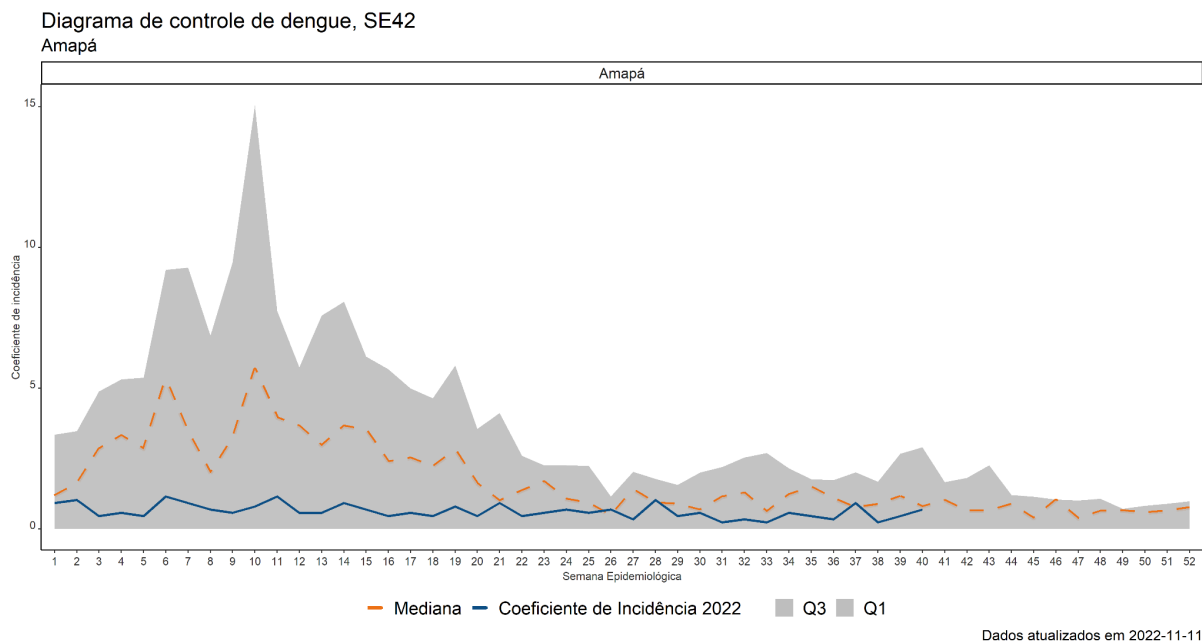
Fonte: Elaborado pelo autor.

Gráfico 3 - Diagrama de Controle de dengue, Alagoas, SE 42 de 2022.



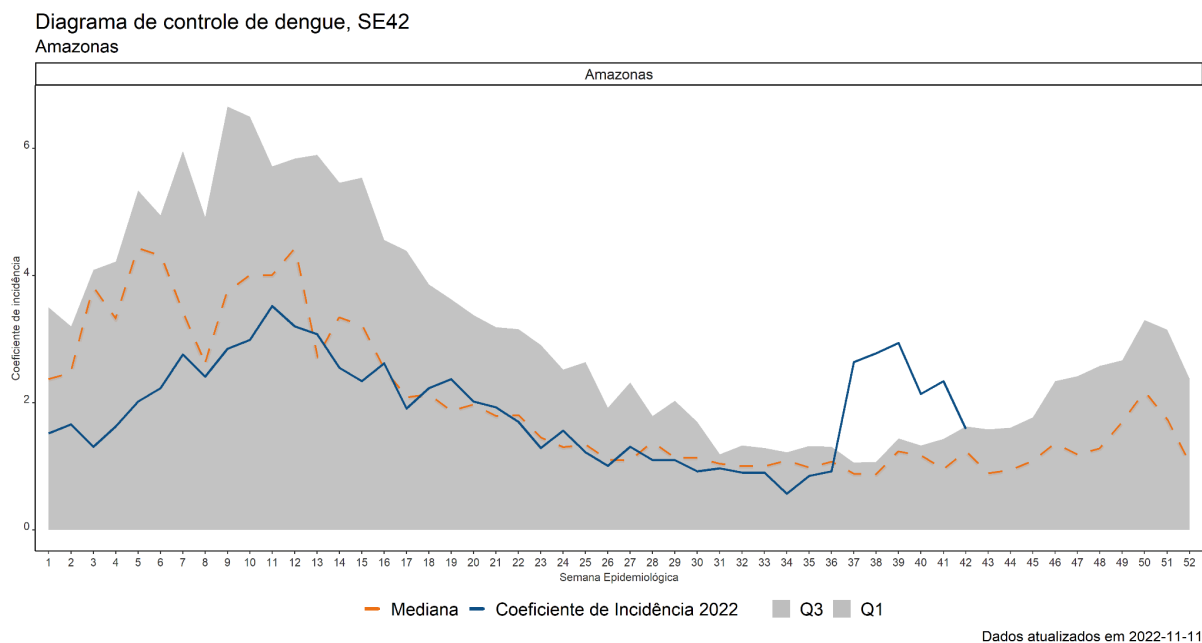
Fonte: Elaborado pelo autor.

Gráfico 4 - Diagrama de Controle de dengue, Amapá, SE 42 de 2022.



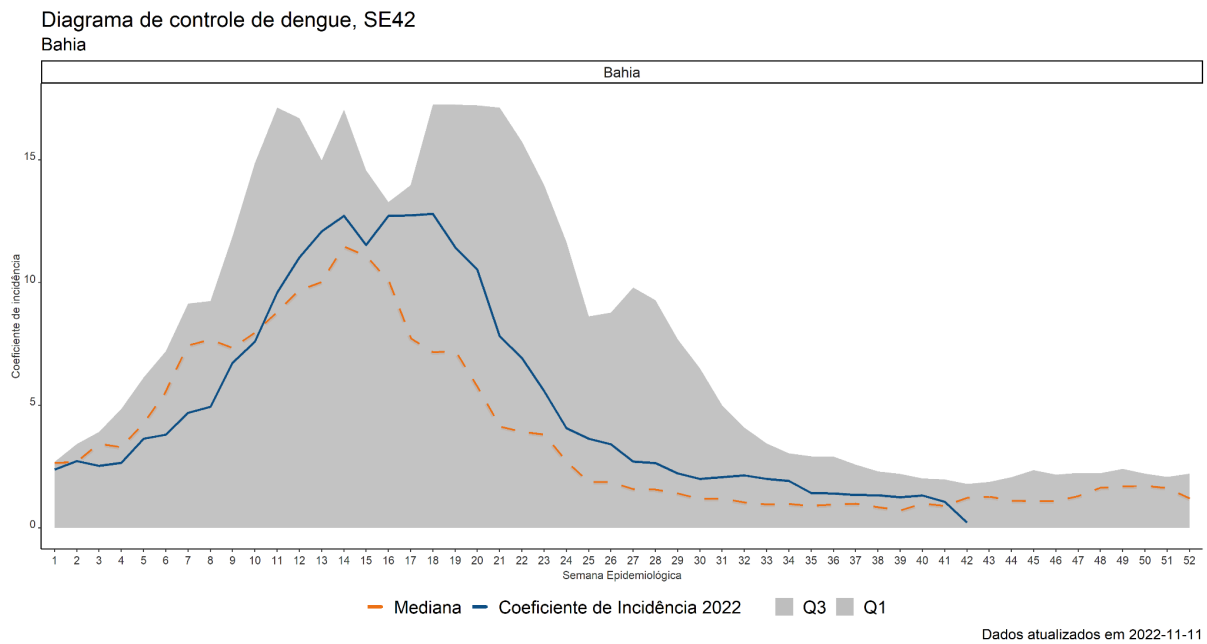
Fonte: Elaborado pelo autor.

Gráfico 5 - Diagrama de Controle de dengue, Amazonas, SE 42 de 2022.



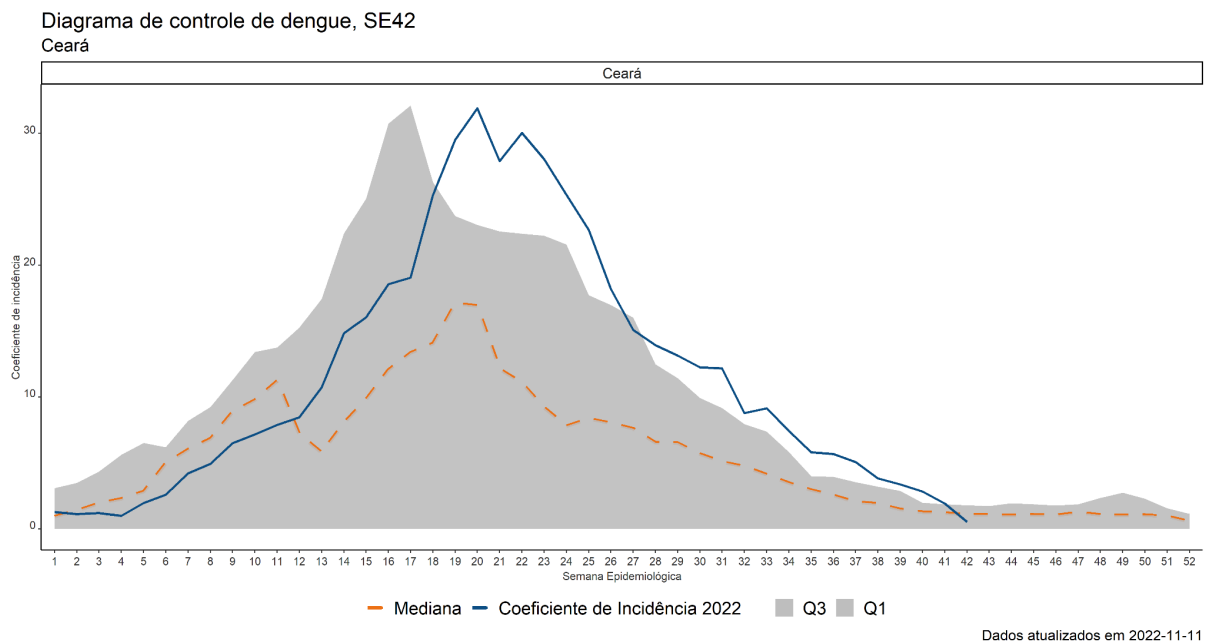
Fonte: Elaborado pelo autor.

Gráfico 6 - Diagrama de Controle de dengue, Bahia, SE 42 de 2022.



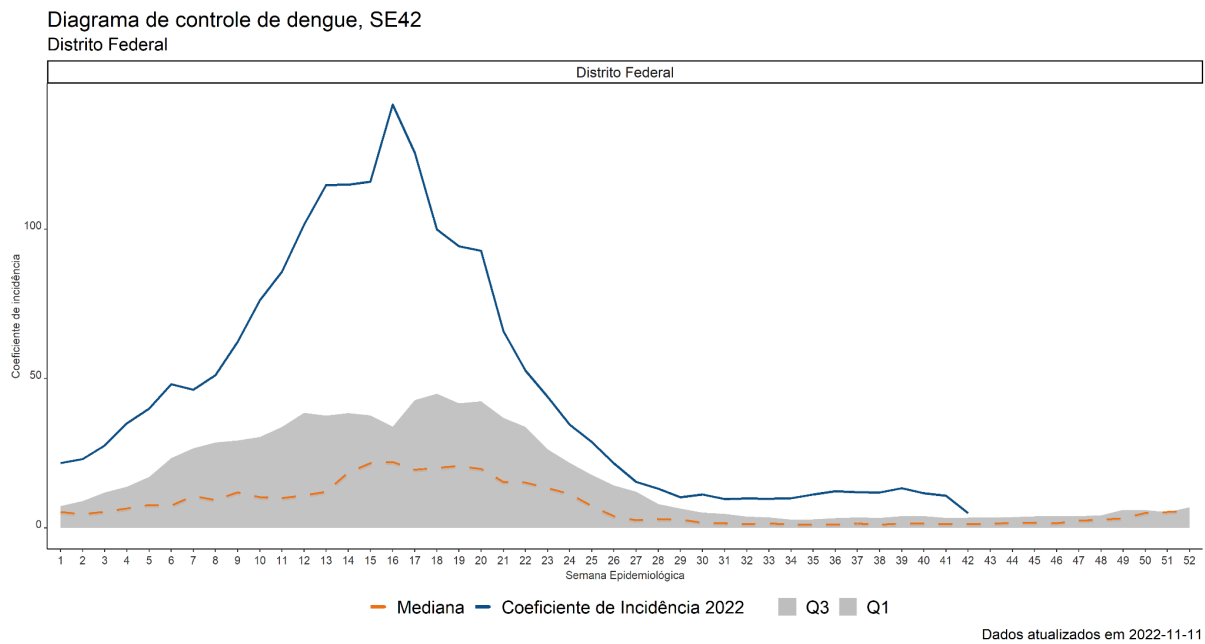
Fonte: Elaborado pelo autor.

Gráfico 7 - Diagrama de Controle de dengue, Ceará, SE 42 de 2022.



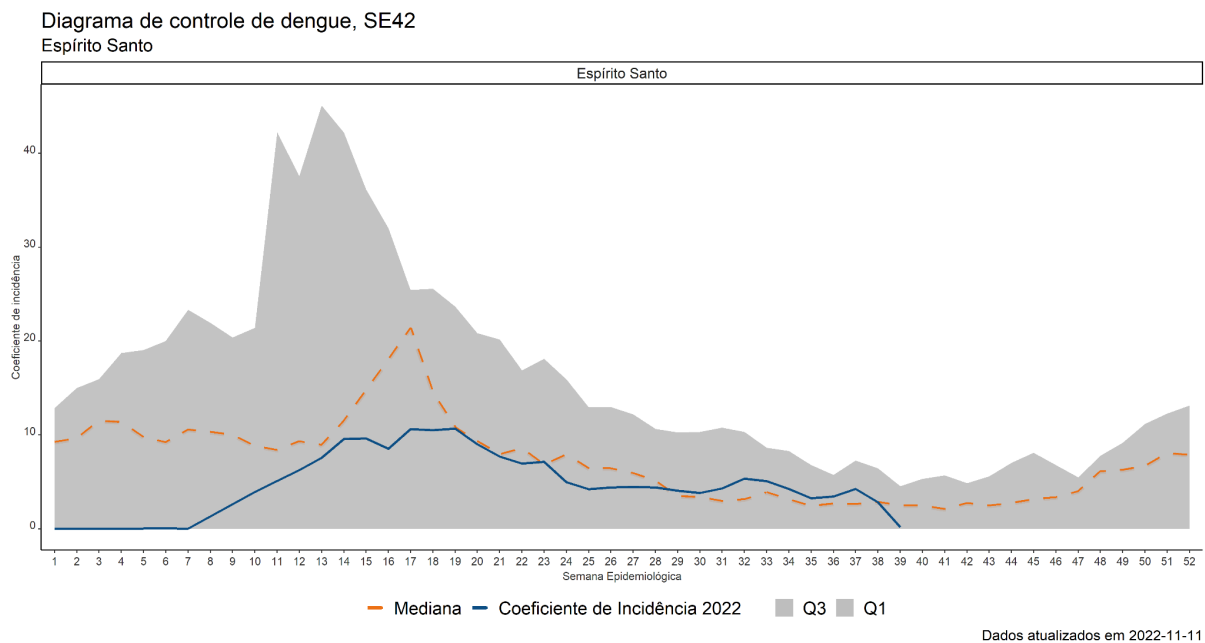
Fonte: Elaborado pelo autor.

Gráfico 8 - Diagrama de Controle de dengue, Distrito-Federal, SE 42 de 2022.



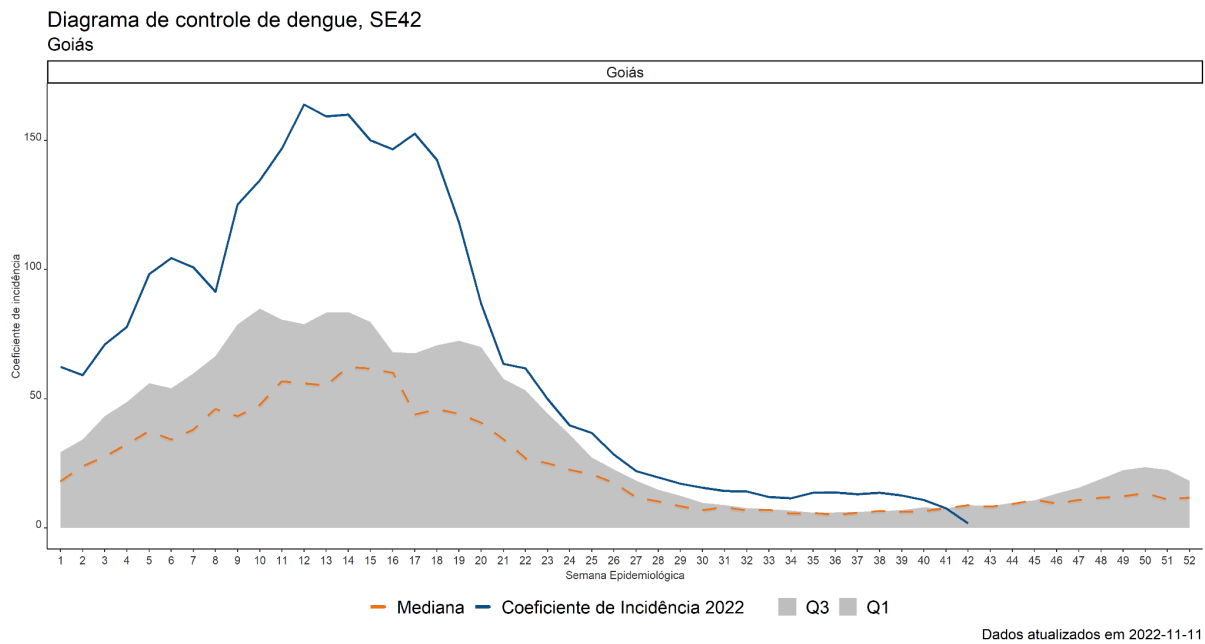
Fonte: Elaborado pelo autor.

Gráfico 9 - Diagrama de Controle de dengue, Espírito Santo, SE 42 de 2022.



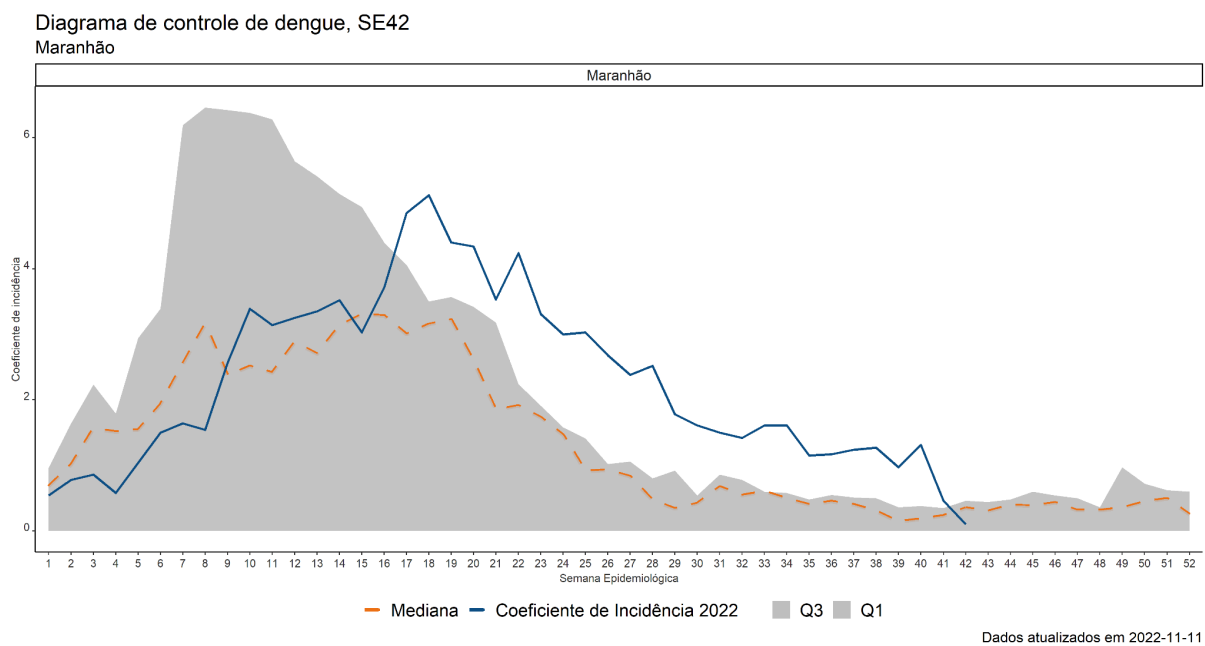
Fonte: Elaborado pelo autor.

Gráfico 10 - Diagrama de Controle de dengue, Goiás, SE 42 de 2022.



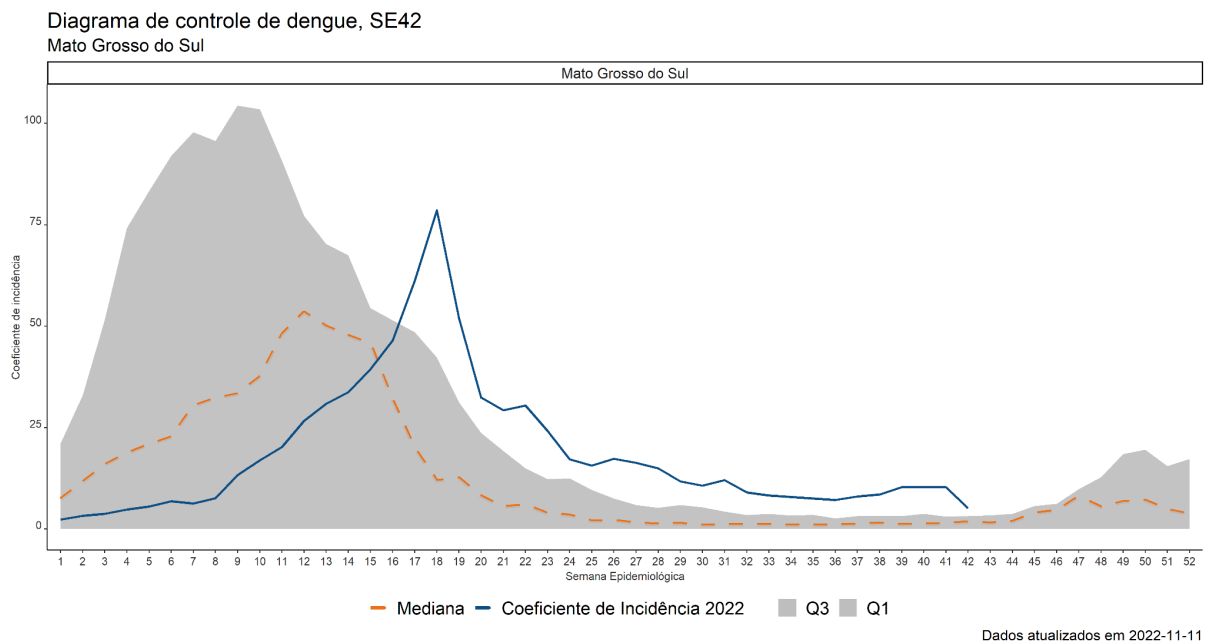
Fonte: Elaborado pelo autor.

Gráfico 11 - Diagrama de Controle de dengue, Maranhão, SE 42 de 2022.



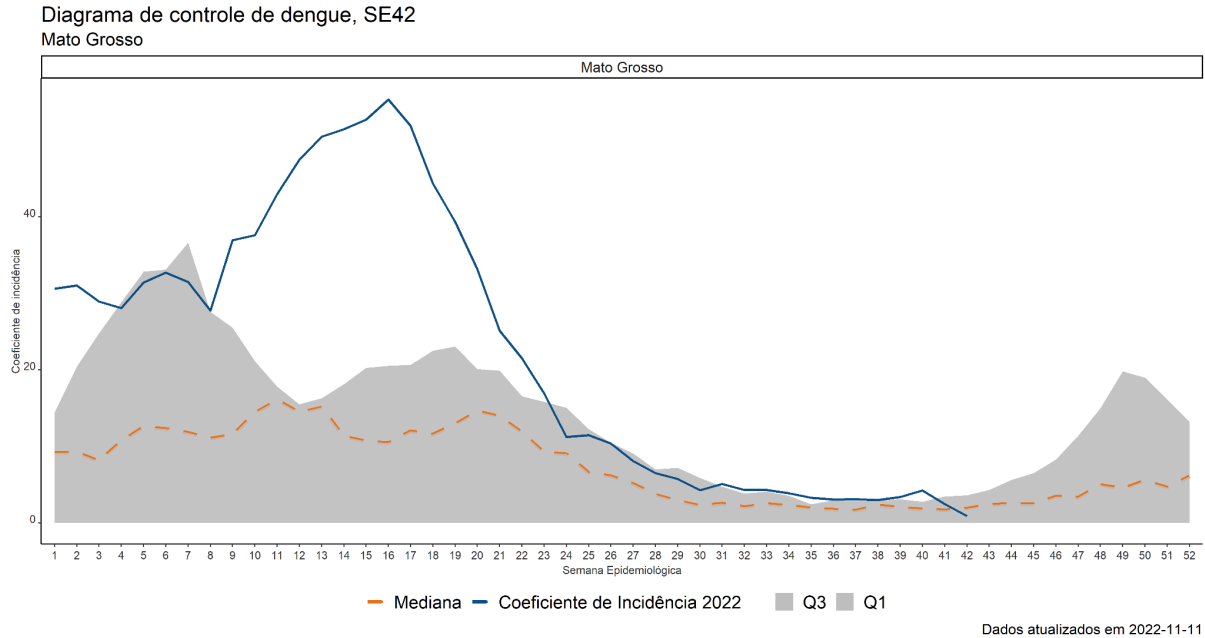
Fonte: Elaborado pelo autor.

Gráfico 12 - Diagrama de Controle de dengue, Mato Grosso do Sul, SE 42 de 2022.



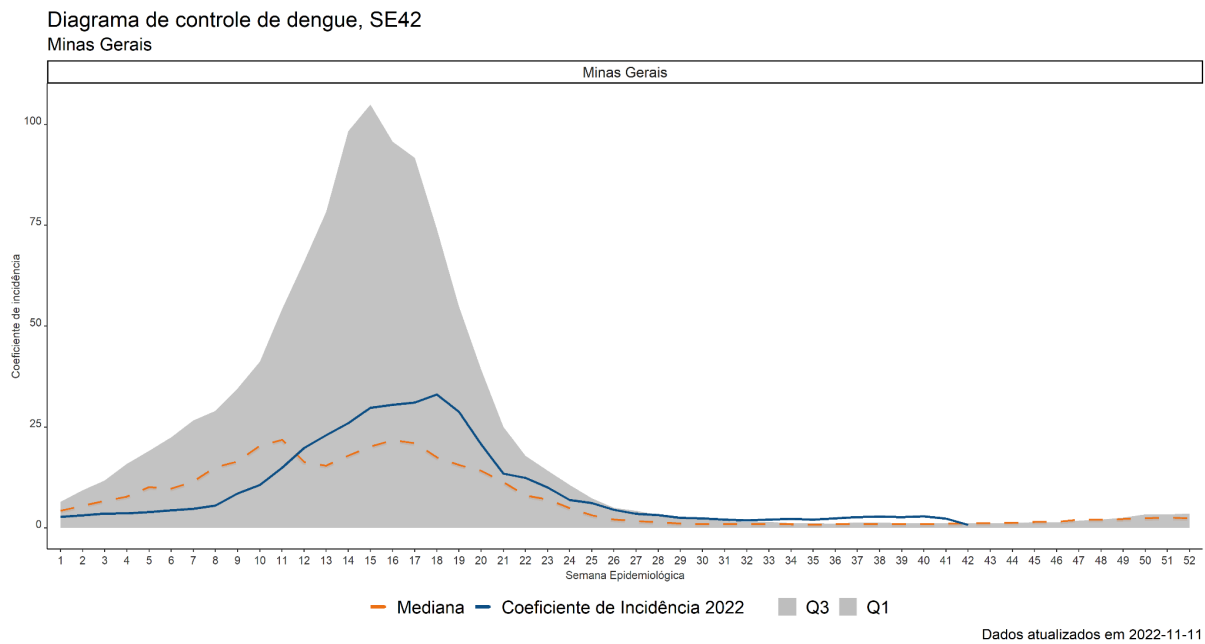
Fonte: Elaborado pelo autor.

Gráfico 13 - Diagrama de Controle de dengue, Mato Grosso, SE 42 de 2022.



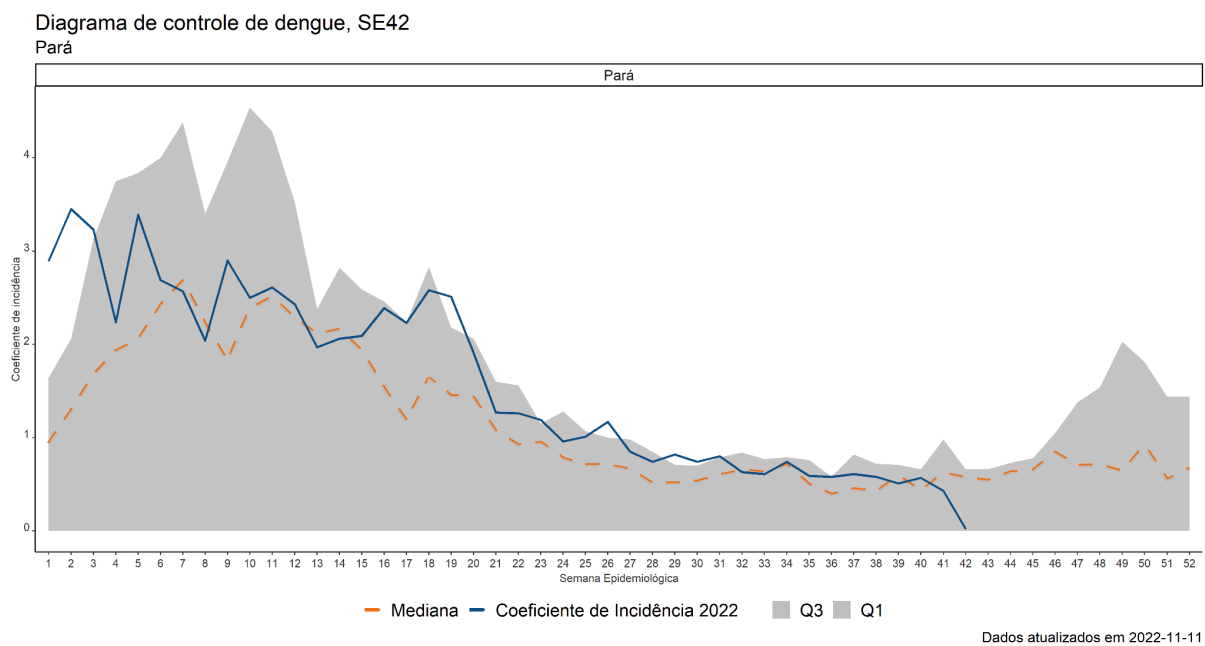
Fonte: Elaborado pelo autor.

Gráfico 14 - Diagrama de Controle de dengue, Minas Gerais, SE 42 de 2022.



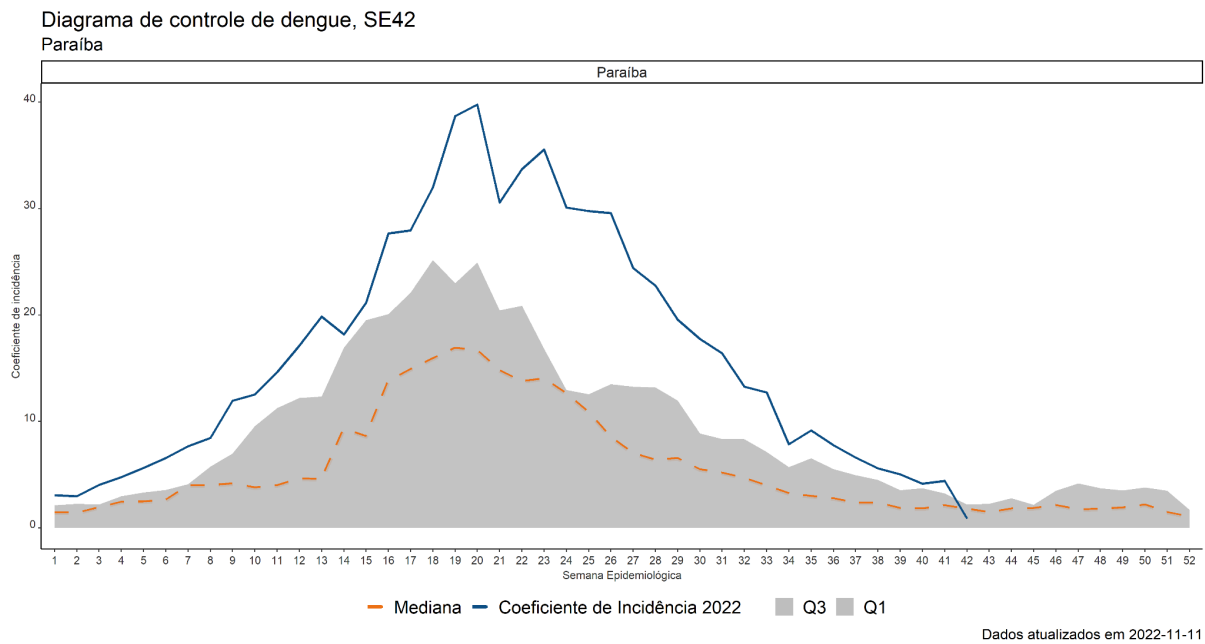
Fonte: Elaborado pelo autor.

Gráfico 15 - Diagrama de Controle de dengue, Pará, SE 42 de 2022.



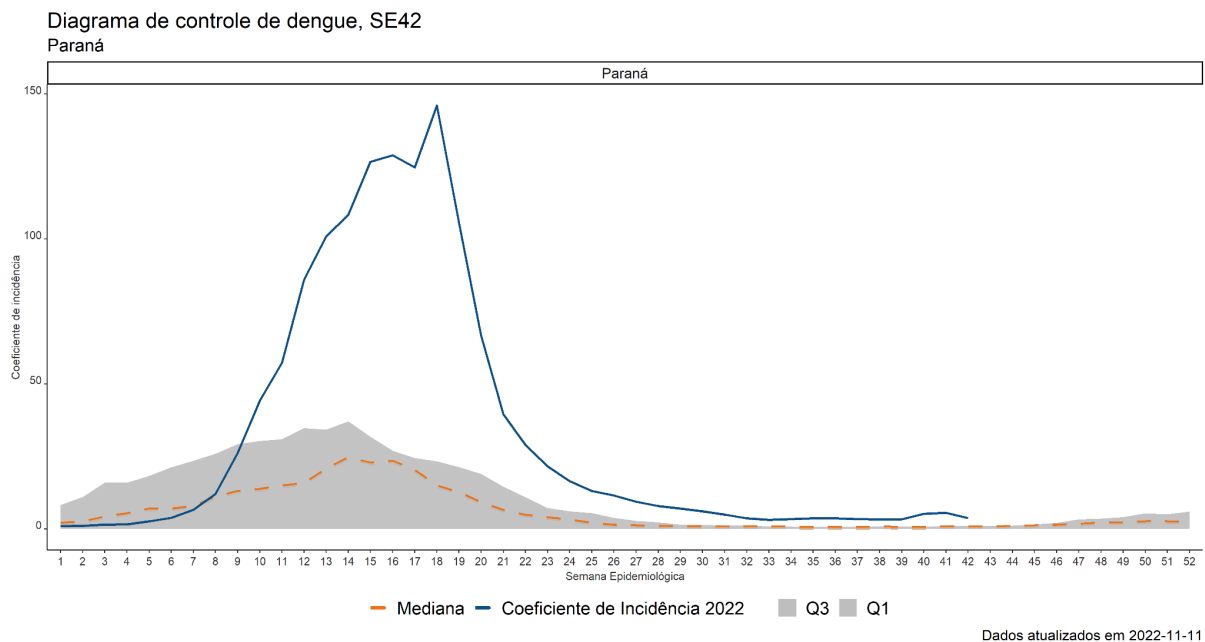
Fonte: Elaborado pelo autor.

Gráfico 16 - Diagrama de Controle de dengue, Paraíba, SE 42 de 2022.



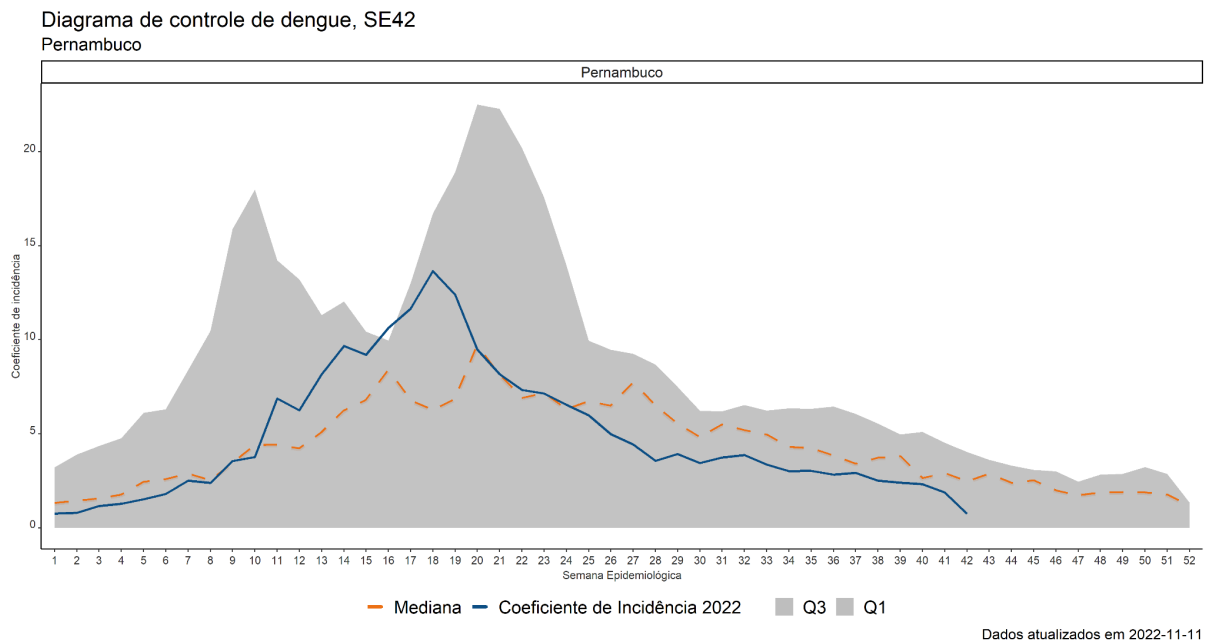
Fonte: Elaborado pelo autor.

Gráfico 17 - Diagrama de Controle de dengue, Paraná, SE 42 de 2022.



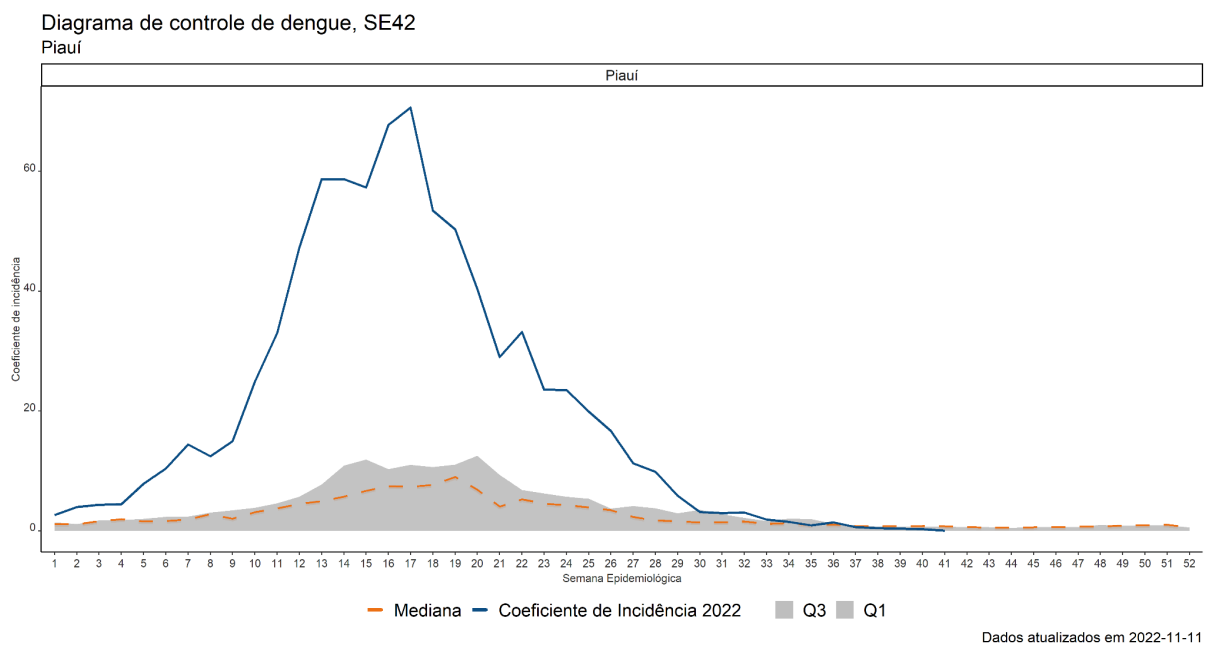
Fonte: Elaborado pelo autor.

Gráfico 18 - Diagrama de Controle de dengue, Pernambuco, SE 42 de 2022.



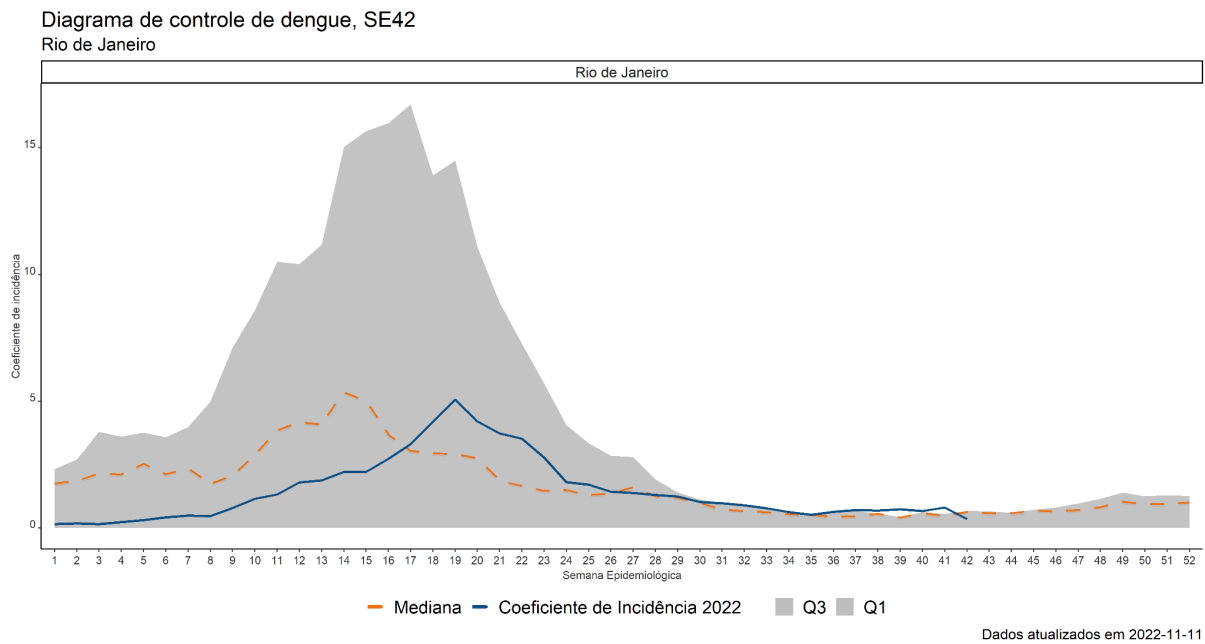
Fonte: Elaborado pelo autor.

Gráfico 19 - Diagrama de Controle de dengue, Piauí, SE 42 de 2022.



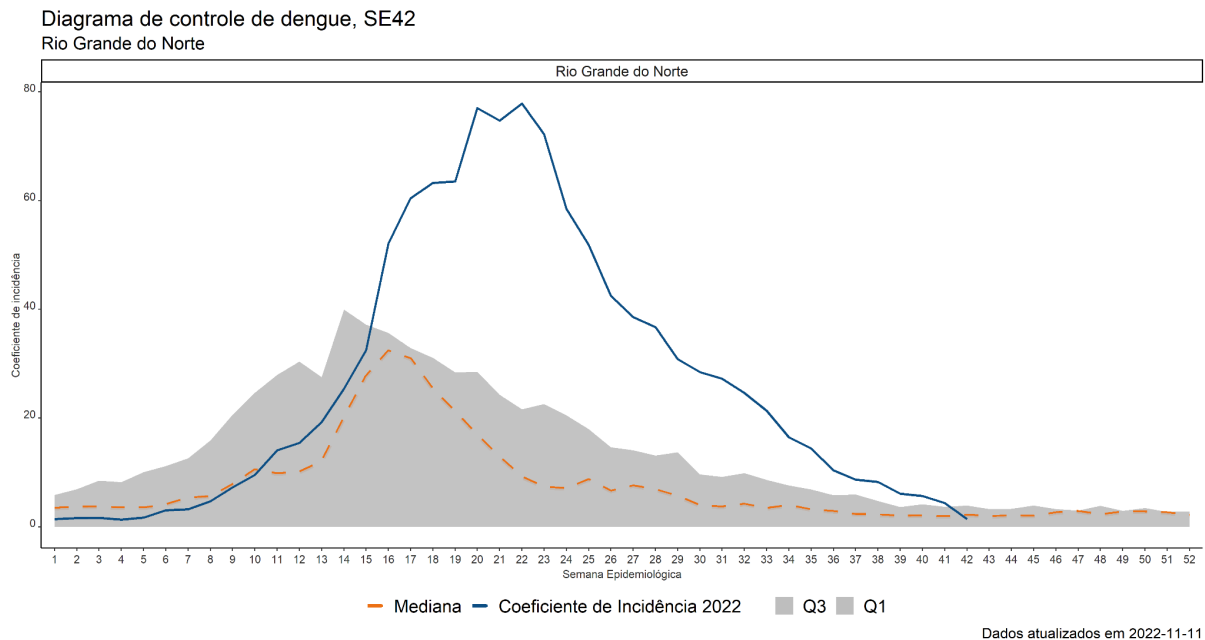
Fonte: Elaborado pelo autor.

Gráfico 20 - Diagrama de Controle de dengue, Rio de Janeiro, SE 42 de 2022.



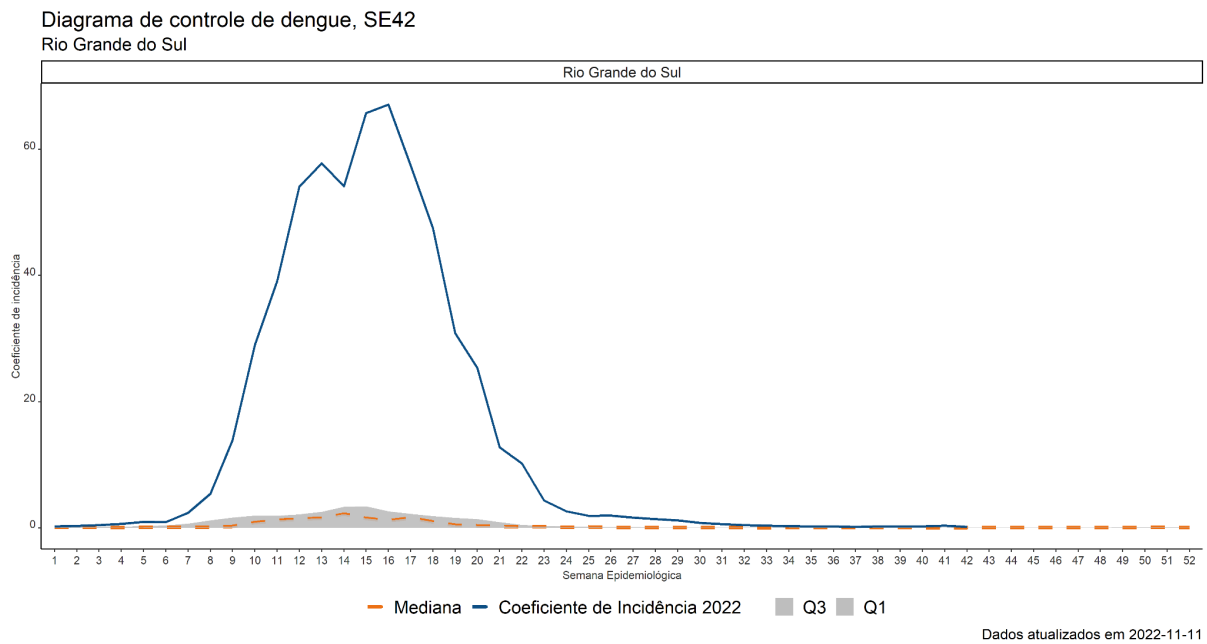
Fonte: Elaborado pelo autor.

Gráfico 21 - Diagrama de Controle de dengue, Rio Grande do Norte, SE 42 de 2022.



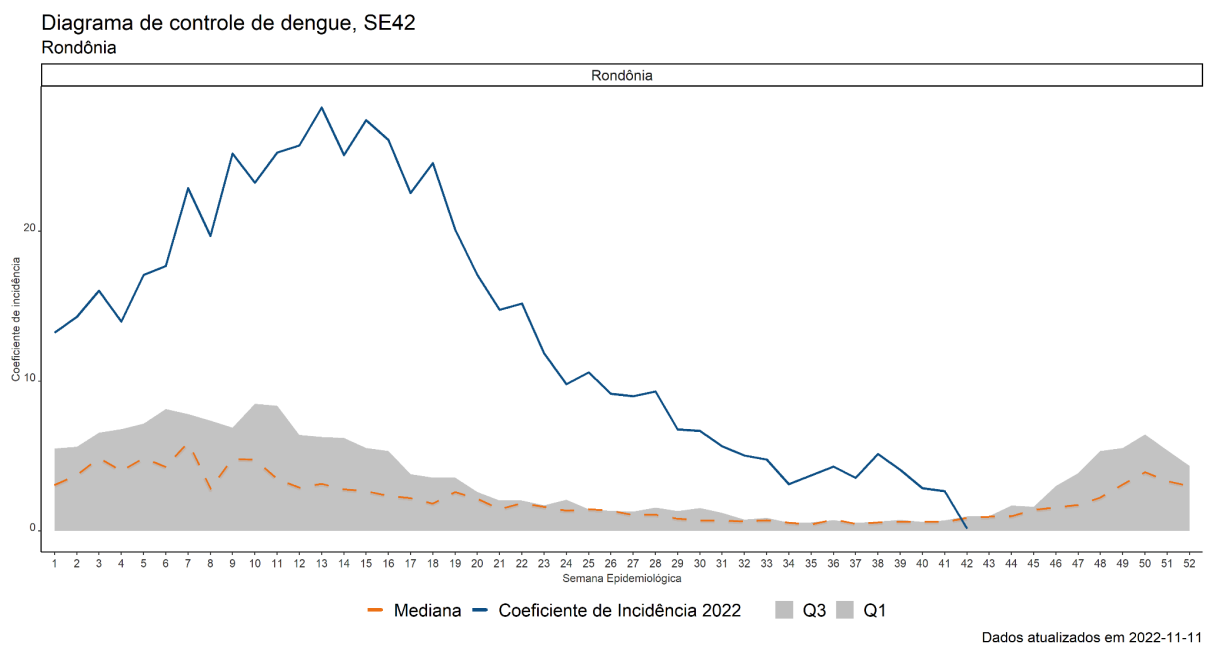
Fonte: Elaborado pelo autor.

Gráfico 22 - Diagrama de Controle de dengue, Rio Grande do Sul, SE 42 de 2022.



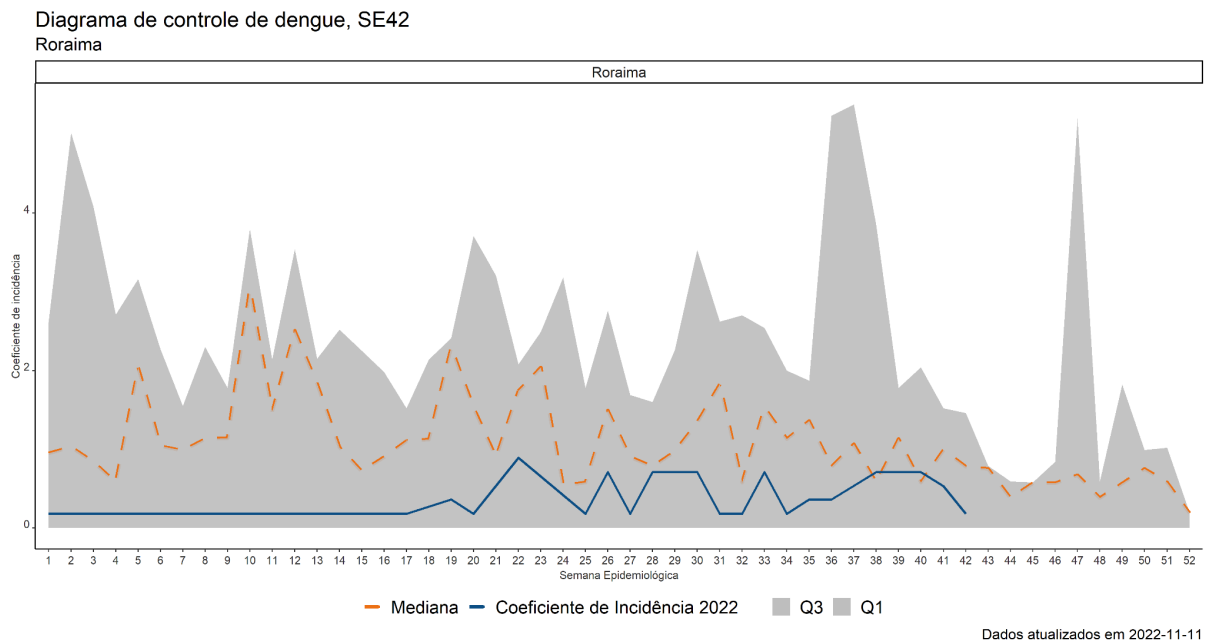
Fonte: Elaborado pelo autor.

Gráfico 23 - Diagrama de Controle de dengue, Rondônia, SE 42 de 2022.



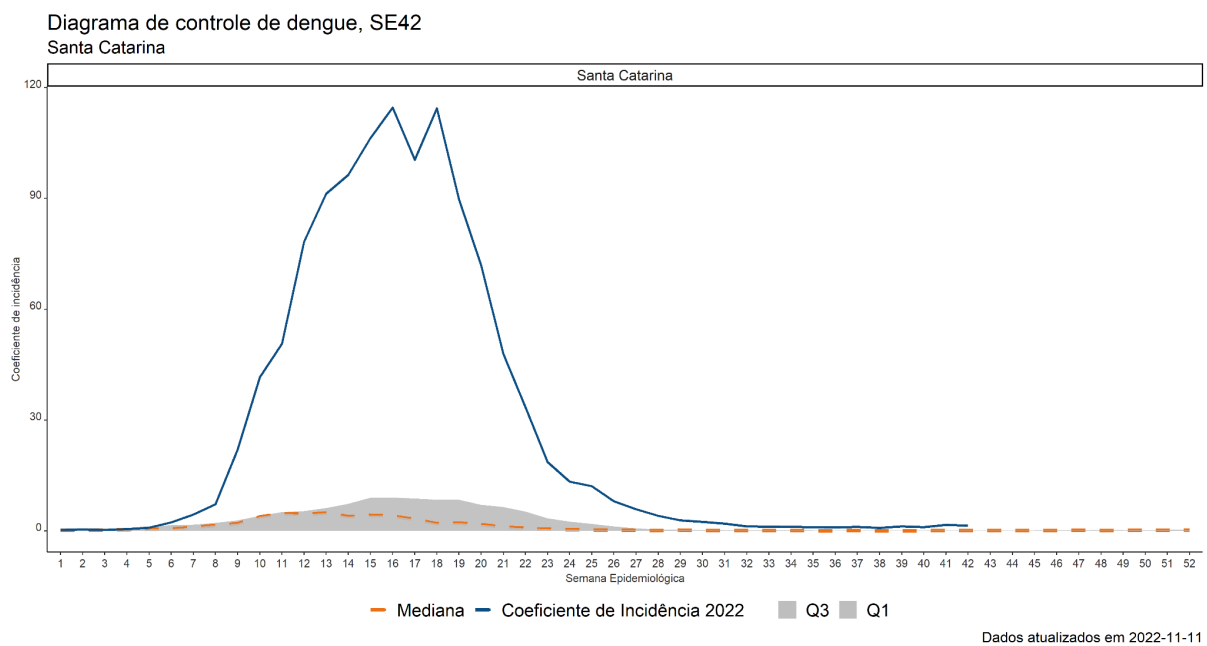
Fonte: Elaborado pelo autor.

Gráfico 24 - Diagrama de Controle de dengue, Roraima, SE 42 de 2022.



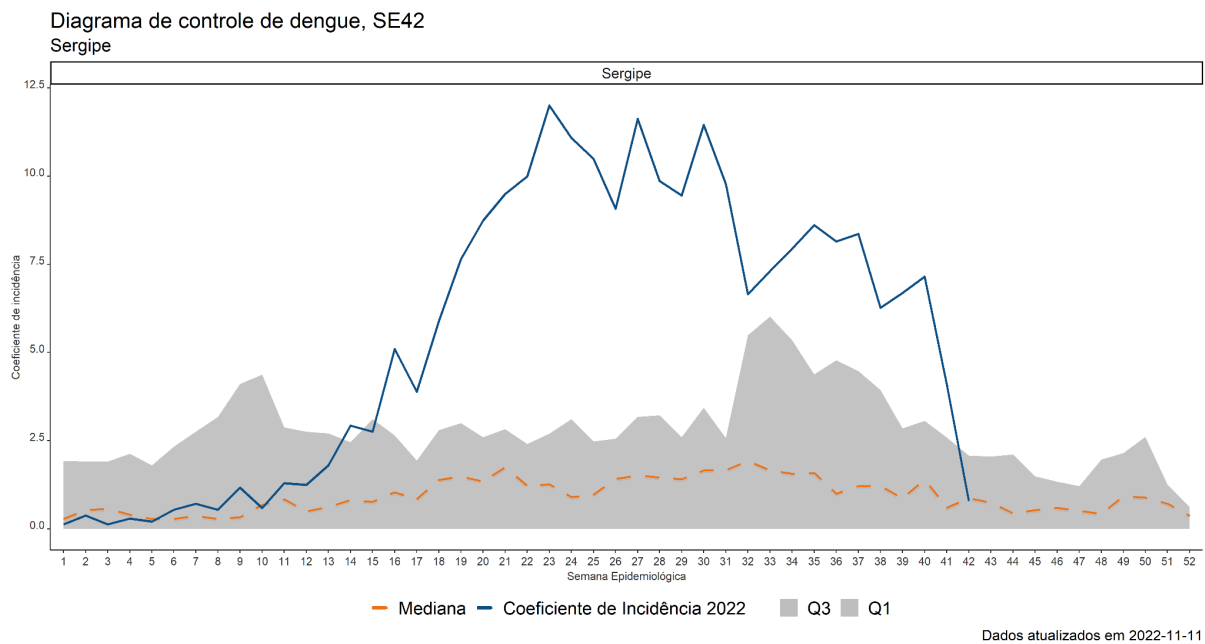
Fonte: Elaborado pelo autor.

Gráfico 25 - Diagrama de Controle de dengue, Santa Catarina, SE 42 de 2022.



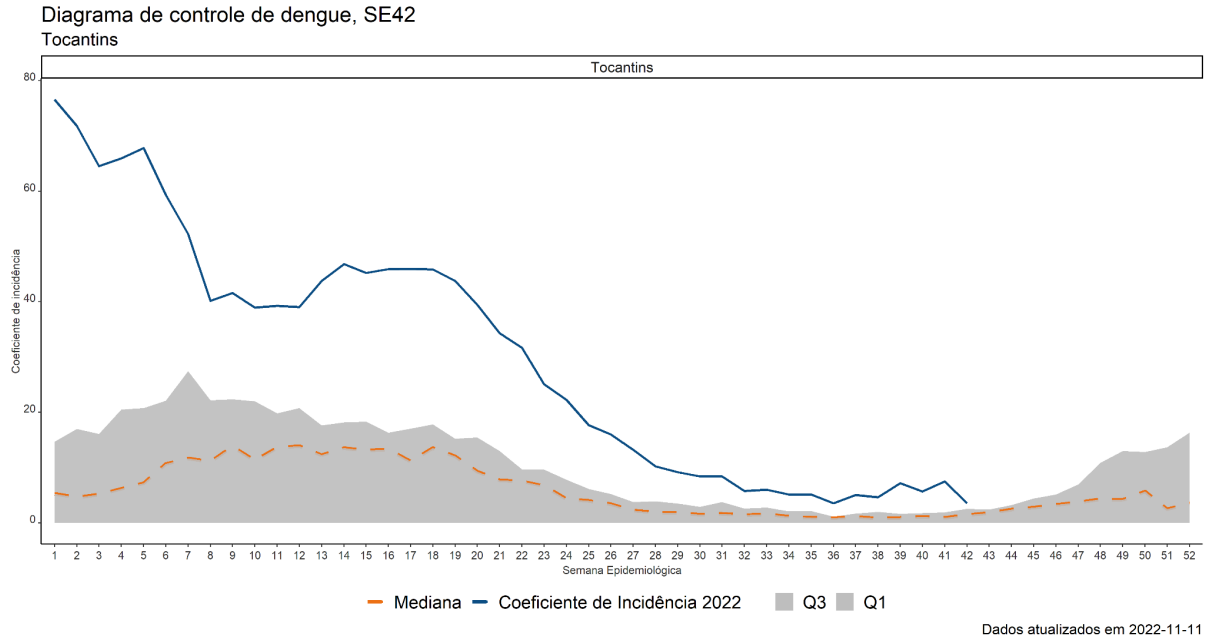
Fonte: Elaborado pelo autor.

Gráfico 26 - Diagrama de Controle de dengue, Sergipe, SE 42 de 2022.



Fonte: Elaborado pelo autor.

Gráfico 27 - Diagrama de Controle de dengue, Tocantins, SE 42 de 2022.



Fonte: Elaborado pelo autor.

APÊNDICE B.

Script para o Diagrama de controle.

```
##### 1° PARTE
```

```
#####
```

```
# _____ Instalar e carregar os pacotes que serão utilizados _____ #
```

```
install.packages("pacman")
```

```
library(pacman)
```

```
pacman::p_load(  
  rio,          # Importa??o e exporta??o  
  tidyverse,   # Manipula??o e tratamento de dados  
  lubridate,  
  dplyr,  
  zoo )        # Manipula??o e tratamento de datas
```

```
if(!require(read.dbc)) install.packages("read.dbc")
```

```
if(!require(foreign)) install.packages("foreign")
```

```
if(!require(hablar)) install.packages("hablar")
```

```
if(!require(stringr)) install.packages("stringr")
```

```
if(!require(readxl)) install.packages("readxl")
```

```
#-----#
```

```
##### 2° PARTE
```

```
#####
```

```
# _____ Ler todos os bancos de dengue de uma vez _____ #
```

```
# _____ OBS: Pode levar um tempo para serem processados _____ #
```

```

x = list.files (path = "C:\\MS\\Dengue\\",
               pattern = "*.dbc")

for(i in 1:length(x)) {
  assign(paste0("dengue", i),
        read.dbc(paste0("C:\\MS\\Dengue\\",
                        x[i])))
}

#-----#

##### 3° PARTE
#####

# Para cada umas das 7 variáveis criadas com a instrução passada,
# vamos seleccionar de cada uma apenas as colunas que iremos utilizar
# e armazenar na mesma variável.

# OBS: Observe q ao executar cada código vai liberando mais espaço na memória
# e reduzindo a quantidade de colunas para 3 .

dengue1 <- dengue1 %>%
  dplyr::select(DT_SIN_PRI,CLASSI_FIN, ID_MN_RESI)

dengue2 <- dengue2 %>%
  dplyr::select(DT_SIN_PRI,CLASSI_FIN, ID_MN_RESI)

dengue3 <- dengue3 %>%
  dplyr::select(DT_SIN_PRI,CLASSI_FIN, ID_MN_RESI)

dengue4 <- dengue4 %>%
  dplyr::select(DT_SIN_PRI,CLASSI_FIN, ID_MN_RESI)

dengue5 <- dengue5 %>%
  dplyr::select(DT_SIN_PRI,CLASSI_FIN, ID_MN_RESI)

dengue6 <- dengue6 %>%
  dplyr::select(DT_SIN_PRI,CLASSI_FIN, ID_MN_RESI)

dengue7 <- dengue7 %>%
  dplyr::select(DT_SIN_PRI,CLASSI_FIN, ID_MN_RESI)

##### 4° PARTE
#####

# _____ Juntar as bases em apenas uma variável _____ #

```

```

dengue15_21 <- rbind(dengue1, dengue2, dengue3,
                    dengue4, dengue5,dengue6, dengue7)
#
____#

# Caso queira exportar as tabelas "unidas" para alguma finalidade:

#export(dengue15_21, "dengue_historico.csv")

#
____#

# A base de 2022 é a mais recente, logo esta em outro formato.

dengue2022 <- read.dbf("C:/MS/Dengue/DENGBR22.dbf")

# Dessa ultima base que baixamos iremos separar apenas as 3 colunas.

dengue2022 <- dengue2022 %>%
  dplyr::select(DT_SIN_PRI,CLASSI_FIN, ID_MN_RESI)

# iremos unir essa base nova com a "antiga".

dengue <- rbind(dengue15_21, dengue2022)

##### 5° PARTE
#####

# para a realização de alguns cálculos precisamos da população,

pop <- read_excel("C:/MS/Pop_UF_2009a2022.xls")

# PIVOTAGEM #

# mudando o formato da tabela para longo
pop <- pop %>%
  #filter(UF == "Brasil") %>%
  #select(- "2022") %>%      # como 2022 não esta completo, não se utiliza.
  pivot_longer(cols = 9:16,  # Escolhendo as colunas que serão utilizadas.

              names_to = "ano", # Definindo o nome da coluna nova.

              values_to = "pop") # Definido o nome da coluna nova , com os
#                               valores respctivos de sua origem.

```

```

#           Verificar a variável de data:           #

class(dengue$DT_SIN_PRI) #se realmente esta no formato de data

min(dengue$DT_SIN_PRI, na.rm = T) # a mínima data nesses dados

max(dengue$DT_SIN_PRI, na.rm = T) # a máxima data nesses dados

#           Filtrar a data no período de 2009 a 2022           #

# A variável dengue vai receber os filtros das datas que forem maiores ou
# iguais a menor data desejada até a maior data desejada.

dengue <- dengue %>%
  #select(DT_SIN_PRI, CLASSI_FIN) %>%
  filter(DT_SIN_PRI >= "2015-01-01") %>%
  filter(DT_SIN_PRI <= "2022-12-31")

# verificar se deu certo os filtros recém feitos

min(dengue$DT_SIN_PRI, na.rm = T)
max(dengue$DT_SIN_PRI, na.rm = T)

# _____
# _____

##### 6° PARTE
#####

# Verificar as categorias da variável dengue

unique(dengue$CLASSI_FIN)

# Excluir os casos descartados

# Neste comando o $ representa que iremos acessar somente a coluna class
# da tabela dengue

dengue$class <- as.character(as.factor(dengue$CLASSI_FIN))

dengue$class <- ifelse(is.na(dengue$class), 99, dengue$class)

dengue <- dengue %>%

```

```

filter(class != 5)

unique(dengue$CLASSI_FIN)

##### 8° PARTE #####

#           Agregar os casos por dia

dengue$uf <- substring(dengue$ID_MN_RESI, 1,2)

dengue2 <- dengue %>%
  dplyr::group_by(DT_SIN_PRI, uf) %>%
  dplyr::summarise(casos = n())%>%
  drop_na(uf)

br <- dengue2 %>%
  group_by(DT_SIN_PRI) %>%
  summarise(casos = sum(casos)) %>%
  mutate(uf = 76) %>%
  select(DT_SIN_PRI, uf, casos)

dengue2$uf <- as.numeric(as.character(dengue2$uf))

dengue2 <- rbind(dengue2, br)

##### 9° PARTE #####

# Transformar a UF em fator e colocar o nome da UF em seu respectivo código

dengue2$ufres <- factor(dengue2$uf,levels=c(11:17,21:29,31:33,35,41:43,50:53, 76),
  labels = c("Rondônia","Acre","Amazonas","Roraima","Pará",
    "Amapá","Tocantins","Maranhão","Piauí","Ceará",
    "Rio Grande do Norte","Paraíba","Pernambuco",
    "Alagoas","Sergipe","Bahia","Minas Gerais",
    "Espírito Santo","Rio de Janeiro","São Paulo",
    "Paraná", "Santa Catarina", "Rio Grande do Sul",
    "Mato Grosso do Sul", "Mato Grosso", "Goiás",
    "Distrito Federal", "Brasil"))

dengue2$ufres <- factor(dengue2$uf,levels=c(11:17,21:29,31:33,35,41:43,50:53),
  labels = c("Rondônia","Acre","Amazonas","Roraima","Pará",
    "Amapá","Tocantins","Maranhão","Piauí","Ceará",
    "Rio Grande do Norte","Paraíba","Pernambuco",
    "Alagoas","Sergipe","Bahia","Minas Gerais",
    "Espírito Santo","Rio de Janeiro","São Paulo",

```

```
"Paraná", "Santa Catarina", "Rio Grande do Sul",  
"Mato Grosso do Sul", "Mato Grosso", "Goiás",  
"Distrito Federal"))
```

```
dengue2 <- dengue2 %>%  
  drop_na(ufres)
```

```
##### 10° PARTE #####
```

```
#Calend?rio  
library(zoo)  
library(lubridate)  
calendario <- seq.Date(from = as.Date("2015-01-01"),  
  to = as.Date("2022-12-31"),  
  by = "days") %>% as.data.frame()
```

```
calendario <- calendario %>%  
  rename("d1_epi_sem_ano"=".") %>%  
  mutate(dia_sem = weekdays(d1_epi_sem_ano)) %>%  
  filter(dia_sem == "domingo")
```

```
calendario <- calendario %>%  
  mutate (epi_sem_ano = paste0(epiweek(d1_epi_sem_ano),  
    "-",  
    epiyear(d1_epi_sem_ano)),  
    epi_sem_ano = ifelse(epiweek(d1_epi_sem_ano)<10,  
    paste0("0", epi_sem_ano),  
    epi_sem_ano)) %>%
```

```
  select(-dia_sem)
```

```
##### 11° PARTE  
#####
```

```
# Criar semana epidemiol?gica na base de dengue e unir com o calend?rio
```

```
dengue2 <- dengue2 %>%  
  mutate( epi_ano    = epiyear(DT_SIN_PRI),  
    epi_semana    = epiweek(DT_SIN_PRI),  
    epi_sem_ano  = paste0(epi_semana, "-", epi_ano),  
    epi_sem_ano  = ifelse(epi_semana<10,  
    paste0("0", epi_sem_ano),  
    epi_sem_ano)) %>%  
  left_join(calendario, by="epi_sem_ano")
```

```
# Agregar os casos por semana e UF
```



```

dengue2 <- dengue2 %>%
  group_by(d1_epi_sem_ano, ufres) %>%
  summarise(casos = sum(casos)) %>%
  drop_na(d1_epi_sem_ano) %>%
  mutate(se = epiweek(d1_epi_sem_ano))

#criar a vari?vel ano na base de dengue

dengue2$ano <- epiyear(dengue2$d1_epi_sem_ano)

pop$ano <- as.numeric(as.character(pop$ano))

#unir base dengue com a pop a partir do ano e ufres

dengue2 <- merge(dengue2, pop, by.x = c("ano", "ufres"),
  by.y = c("ano", "UF"), all.x = T)

##### 12° PARTE
#####

#          coeficiente de incid?ncia

dengue2$ci <- round(dengue2$casos/dengue2$pop * 100000, 2)

#          Filtrar anos 2019, 2020, 2021 e 2022

dengue2_1 <- dengue2 %>%
  filter(d1_epi_sem_ano >= "2019-01-01") %>%
  filter(d1_epi_sem_ano <= "2022-12-31")

### Histograma por UF usando um loop

#
# UF <- unique(dengue2$ufres)
#
# UF <- unique(dengue2_1$ufres)
#
# plotUF =list()
# for(i in 1:length(UF)){
#
#   plotUF[[i]]=ggplot()+
#     geom_histogram(data=dengue2_1[which(dengue2_1$ufres == UF[i]),],
#     aes(x = d1_epi_sem_ano, y = casos),
#     stat = "identity",
#     fill="#064c72",

```

```

#         color="#0C4E84") +
#
#   labs(title = "Casos de dengue, SE42",
#         subtitle = UF[i],
#         x   = "Data de início dos sintomas",
#         y   = "Nº de casos") +
#   theme_classic()
#
#   ggsave(plotUF[[i]],
#         filename = paste("H:\\MS\\DENGUE_SE42_24.10.2022\\",
#                           UF[i], "_histograma.png", sep=""),
#         height = 8, width = 15, dpi = 320)
# }

# curva epidêmica por UF usando um loop

#
# for(i in 1:length(UF)){
#
#   plotUF[[i]]=ggplot()+
#     geom_line(data=dengue2_1[which(dengue2_1$ufres == UF[i]).],
#               aes(x=as.factor(se),
#                   y=ci, group=as.factor(ano),
#                   color = as.factor(ano)),
#               size = 1,
#               linetype = "solid") +
#     theme_classic()+
#     scale_colour_manual(values = c("#FF8C00", "#336633", "#3300CC", "#CC0000")) +
#     theme(legend.title = element_blank(),
#           legend.position = "bottom")+
#     labs(title   = "Curva epidêmica de dengue, SE42",
#          subtitle = UF[i],
#          x       = "Semana Epidemiológica",
#          y       = "Coeficiente de incidência") +
#     theme(legend.position = "bottom", legend.box = "vertical")+
#     theme(axis.text.x = element_text(color = "grey20", size = 10, angle = 0,
#                                       hjust = .5, vjust = .5, face = "plain"),
#           axis.text.y = element_text(color = "grey20", size = 10, angle = 0,
#                                       hjust = 1, vjust = 0, face = "plain"),
#           axis.title.x = element_text(color = "grey20", size = 10, angle = 0,
#                                       hjust = .5, vjust = 0, face = "plain"),
#           axis.title.y = element_text(color = "grey20", size = 10, angle = 90,
#                                       hjust = .5, vjust = .5, face = "plain"),
#           legend.text=element_text(size=10),
#           text=element_text(size=10),
#           title = element_text(size=10))
# }

```

```

#
# ggsave(plotUF[[i]],filename=paste("H:\\MS\\DENGUE_SE42_24.10.2022\\",
#                                     UF[i], "_ce.png", sep=""),
#       height = 8, width = 15, dpi = 320)
# }

# C?lculo das estat?sticas descritivas (mediana e quartis)

dengue12_21 <- dengue2 %>%
  filter(ano != 2022) %>%
  mutate(se = epiweek(d1_epi_sem_ano))%>%
  group_by(se, ufres) %>%
  summarise(mediana = median(ci), #mediana
            Q1 = quantile(ci, probs = c(0.25)),
            Q3 = quantile(ci, probs = c(0.75)))%>% #quartis
  mutate(low = round(Q1, 2), high = round(Q3, 2))

# Constru??o da semana epidemiol?gica para o ano de 2022

dengue2022_ <- dengue2 %>%
  filter(ano == 2022) %>%
  select(ci, se, ufres)

dengue12_21 <- dengue12_21 %>%
  filter(se != 53)

dengue12_21$se <- factor(dengue12_21$se, levels = c(1:52))

#### diagrama por UF usando loop
library(ggplot2)
UF <- unique(dengue2$ufres)

plotDiagUF =list()
for(i in 1:length(UF)){

  plotDiagUF[[i]]= ggplot()+
    geom_ribbon(data = dengue12_21[which(dengue12_21$ufres==UF[i]),],
aes(ymin=mediana, ymax=high, x = se,group =1, fill = "Q3"), alpha=0.9)+
    geom_line(data = dengue12_21[which(dengue12_21$ufres==UF[i]),], aes(x = se, y=
mediana, group =1, colour = "Mediana"),linetype = "dashed", size = 1.5)+
    geom_ribbon(data = dengue12_21[which(dengue12_21$ufres==UF[i]),], aes(ymin=0,
ymax=mediana, x = se,group =1, fill = "Q1"), alpha=0.9)+
    #geom_line(data = dengue09_21, aes(x = se, y= low, group =1, colour = "Q1"), size = 1)+
    #geom_line(data = dengue09_21, aes(x = se, y= high,group =1, colour = "Q3"), size = 1)+

```

```

geom_line(data = dengue2022_[which(dengue2022_$ufres==UF[i]),], aes(x = se, y= ci,
group =1,colour = "Coeficiente de Incidência 2022"), size = 1, linetype = "solid")+
scale_color_manual("", values = c("Mediana" = "#ec7014","Coeficiente de Incidência
2022" = "#0C4E84"))+
scale_fill_manual("", values = c("Q3" = "#bdbdbd", "Q1" = "#bdbdbd"))+
facet_wrap(~ufres)+
theme_classic()+
theme(axis.text.x = element_text(color = "grey20", size = 10, angle = 0, hjust = .5, vjust =
.5, face = "plain"),
axis.text.y = element_text(color = "grey20", size = 10, angle = 0, hjust = 1, vjust = 0,
face = "plain"),
axis.title.x = element_text(color = "grey20", size = 10, angle = 0, hjust = .5, vjust = 0,
face = "plain"),
axis.title.y = element_text(color = "grey20", size = 10, angle = 90, hjust = .5, vjust = .5,
face = "plain"),
legend.text=element_text(size=16),
text=element_text(size=16),
title = element_text(size=16))+
labs(title = "Diagrama de controle de dengue, SE42",
subtitle = UF[i],
x = "Semana Epidemiológica",
y = "Coeficiente de incidência",
caption = str_glue("Dados atualizados em {today()}")) +
theme(legend.position = "bottom",
legend.box = "horizontal")

ggsave(plotDiagUF[[i]],filename=paste("C:\\MS\\Dengue\\",UF[i],"_diagrama.png",sep=""),
height = 8, width = 15, dpi = 320)
}

```

APÊNDICE C.

Script para a tabela que gera o alerta de surto.

```
# Vou precisar de duas tabelas para fazer o join(ou merge()), a q contem o ci  
# é a(dengue2022_) a outra contem a mediana, e o Q3 (dengue12_21).
```

```
##### 1º PASSO  
#####
```

```
# SEPARAR AS COLUNAS QUE QUER UTILIZAR
```

```
tabela <- dengue12_21 %>%  
  dplyr::select(ufres,mediana, Q3)
```

```
tabela_1 <- tabela
```

```
# verificar o tipo de cada coluna da base
```

```
glimpse(tabela_1)  
glimpse(dengue2022_)
```

```
##### PARTE 2  
#####
```

```
# verificar somente uma unica coluna da variavel
```

```
class(tabela_1$se)  
class(dengue2022_$se)
```

```
# Verificando o tipo de variável na coluna
```

```
# COMO SÃO DIFERENTES, PADRONIZAR.
```

```
tabela_1$se <- as.integer(tabela_1$se)  
dengue2022_$se <- as.integer(dengue2022_$se)
```

```

# unir a coluna se + ufres

unir_se_uf_tabela_1 <- tabela_1 %>%
  mutate(
    SE_UF = str_c(se,"_",ufres)
  )

unir_se_uf_dengue <- dengue2022_ %>%
  mutate(
    SE_UF = str_c(se,"_",ufres)
  )

# pronto, ja pode usar o join.

#
# teste <- left_join(
#
# # Unindo a tabela {"tabela_1"}
# x = unir_se_uf_tabela_1,
#
#
# # com a tabela {"dengue2022_"}
# y = unir_se_uf_dengue,
#
# # Selecionando as colunas `SE_UF` e `SE_UF` para unir os bancos de
# # dados
# by = c("SE_UF" = "SE_UF")
# )
#

teste_2 <- merge(unir_se_uf_tabela_1,
  unir_se_uf_dengue,
  by = "SE_UF",
  all = TRUE)

# VAI DIZER SE A PREVALÊNCIA É MAIOR QUE A MEDIANA

teste_3 <- teste_2 %>%
  mutate(      # vai criar uma coluna chamada RISCO_mediana

```

```

RISCO_mediana = ifelse(ci > mediana,
                        TRUE,
                        FALSE)
)

# VAI DIZER SE A PREVALÊNCIA É MAIOR QUE O Q3
teste_4 <- teste_3 %>%
  mutate(              # vai criar uma coluna chamada RISCO_Q3
    RISCO_Q3 = ifelse(ci > Q3,
                      TRUE,
                      FALSE)
  )

# Selecionar as colunas essenciais

teste_5 <- teste_4 %>%
  dplyr::select(se.x,
               ufres.x,
               mediana,
               Q3,
               ci,
               RISCO_mediana,
               RISCO_Q3)

# Validar se até esse momento a análise esta correta " curadoria". O método de
verificação
# para saber se o andamento do código esta correto
# auxiliando na rotina do profissional de saúde,

# função filter() para analisar por semana

semana_1 <- filter(teste_5, se.x == "1")

semana_2 <- filter(teste_5, se.x == "2")

semana_3 <- filter(teste_5, se.x == "3")

semana_4 <- filter(teste_5, se.x == "4")

semana_5 <- filter(teste_5, se.x == "5")

```

```
semana_6 <- filter(teste_5, se.x == "6")

semana_42 <- filter(teste_5, se.x == "42")

# filtro em ordem crescente

filtro_SP <- filter(teste_5, ufres.x == "São Paulo")
filtro_SP <- arrange(filtro_SP,se.x)

export(filtro_SP,"Tabela_SP.xlsx") #exportar tabela em excel no WD

semana_1 <- arrange(semana_1, RISCO_mediana)
```