


RtabnetSP: pacote R para extração de indicadores de saúde do estado de São Paulo*

doi: 10.1590/S1679-4974202100020

RtabnetSP: an R package for retrieving São Paulo State health status indicators, Brazil

RtabnetSP: paquete R para extracción de indicadores de salud del estado de São Paulo, Brasil

João Henrique de Araújo Morais¹ –  orcid.org/0000-0003-3258-1498

Thais Cláudia Roma de Oliveira Konstantyner² –  orcid.org/0000-0002-2289-1866

Álvaro Luiz Fazenda³ –  orcid.org/0000-0002-4052-1113

Arnaldo Sala⁴ –  orcid.org/0000-0002-8941-8045

Camila Bertini Martins² –  orcid.org/0000-0002-8252-8815

¹Fundação Oswaldo Cruz, Escola Nacional de Saúde Pública Sergio Arouca, Rio de Janeiro, RJ, Brasil

²Universidade Federal de São Paulo, Departamento de Medicina Preventiva, São Paulo, SP, Brasil

³Universidade Federal de São Paulo, Instituto de Ciência e Tecnologia, São José dos Campos, SP, Brasil

⁴Secretaria de Estado da Saúde, Coordenadoria de Regiões de Saúde, São Paulo, SP, Brasil

Resumo

Indicadores de saúde representam uma importante ferramenta de acompanhamento de desempenho de ações em Saúde Pública, permitindo a avaliação de intervenções realizadas, bem como a identificação de tendências e regiões prioritárias para alocação de recursos. Com o objetivo de aumentar a praticidade nas tarefas de análise e manipulação de dados desses indicadores, foi criado um pacote R. O pacote rtabnetSP realiza requisições aos servidores TabNet da página eletrônica da Secretaria de Estado da Saúde de São Paulo, recuperando e tratando tais dados para utilização do usuário. Este artigo apresenta o pacote rtabnetSP e suas funções, modo de instalação e uso; traz também exemplos de suas funcionalidades, que permitem a visualização, busca e seleção, entre uma lista de indicadores, do conteúdo desejado, além da obtenção dos dados agregados pelo nível de regionalização disponível na matriz de dados, conferindo maior agilidade a tarefas de gestão em saúde do estado de São Paulo.

Palavras-chave: Software; Processamento Eletrônico de Dados; Indicadores Básicos de Saúde; Sistemas de Informação em Saúde; Gestão em Saúde.

*Artigo derivado de monografia de conclusão de curso intitulada 'Desenvolvimento de aplicativo R/Shiny para auxílio na visualização e análise de indicadores de saúde no estado de São Paulo', apresentada por João Henrique de Araújo Morais junto ao Curso de Ciência da Computação da Universidade Federal de São Paulo, em 2019.

Endereço para correspondência:

João Henrique de Araújo Morais – Escola Nacional de Saúde Pública Sergio Arouca, Fundação Oswaldo Cruz, Rua Leopoldo Bulhões, nº 1480, Manguinhos, Rio de Janeiro, RJ, Brasil. CEP: 21041-210
E-mail: joao.tlp@gmail.com

Introdução

Desde a implantação do Sistema Único de Saúde (SUS) brasileiro, a análise sistemática de dados e informações em saúde é tida como fundamental para a avaliação de seu desempenho. Esse processo auxilia a tomada de decisões e ações relativas ao sistema, como a programação e o planejamento de serviços, alocação e distribuição de recursos e avaliação do impacto de intervenções realizadas.^{1,2}

O pacote rtabnet é capaz de acessar os indicadores de saúde disponibilizados na plataforma TabNet da página da Secretaria de Estado da Saúde de São Paulo.

Dada a importância do tema, a expansão do acesso a dados e informações estruturadas tem sido prioridade no planejamento de sistemas de informações em saúde no nível global³ e, por conseguinte, uma das principais linhas de atuação do Departamento de Informática do SUS (Datasus). O Datasus, atualmente, disponibiliza dados provenientes dos sistemas de informações do Ministério da Saúde em interfaces, a exemplo do TabDOS (versão para DOS), TabWin (versão para Windows), TabNet (versão para a internet) e TabSQL (versão para banco de dados Oracle).⁴ Estas ferramentas permitem ao usuário a manipulação sobre os arquivos de banco de dados de tipo DBF, extensão-padrão dos arquivos tabuláveis das bases de dados onde se encontram os registros das notificações, nos diversos sistemas de informações, possibilitando a realização de tabulações e cálculo de indicadores de saúde pelo interessado, nos diferentes níveis de agregação, como municipal, estadual ou federal.

Observam-se, contudo, limitações inerentes aos softwares disponibilizados, como restrições do sistema operacional e dificuldades de manipulação, visualização e análise dos dados pela plataforma. Uma ferramenta possível de estender essas funcionalidades é o software estatístico R, disponibilizado na forma de código aberto, apropriado para análises diversas previstas.

Existem algumas ferramentas desenvolvidas em linguagem R que recuperam dados dos sistemas de informações do SUS na forma de microdados, como o microdatasus.⁵ Porém, quando se trata de dados agregados, como indicadores de saúde, não há uma ferramenta específica além da extração manual do TabNet ou do cálculo a partir dos microdados.

No contexto descrito, decidiu-se desenvolver um pacote para o software R, denominado rtabnet, capaz de acessar os indicadores de saúde disponibilizados na plataforma TabNet da página da Secretaria de Estado da Saúde de São Paulo (SES/SP) e realizar o download dos dados de indicadores já calculados, no período, região e nível de agregação desejados.

Este artigo teve por objetivo apresentar o pacote rtabnet e suas funções, modo de instalação e uso; traz também exemplos de suas funcionalidades, que permitem, entre uma lista de indicadores, a visualização, busca e seleção do conteúdo de interesse, de maneira a se obterem os resultados desejados e, assim, se conferir maior agilidade a tarefas de gestão em saúde no estado de São Paulo.

Métodos

O pacote rtabnet foi desenvolvido em linguagem R para versões superiores à 3.4.4. O R é uma linguagem de programação estatística de excelência, desenvolvida como um software de acesso livre,⁶ para manipulação, análise e visualização gráfica de dados, além de ser expansível mediante utilização de 'pacotes', ou seja, bibliotecas para organizar e padronizar funções extras do R, cada uma desenvolvida com uma finalidade específica. Além disso, vem se tornando frequente a utilização de pacotes R como ferramentas para obtenção ou manejo de dados relacionados à Saúde Pública.^{5,7,8}

Para esta apresentação, foi considerada a matriz de indicadores do Projeto de Avaliação de Desempenho de Sistemas de Saúde (PROADESS),⁹ haja vista os indicadores de saúde serem fundamentais para análises socioespaciais da situação de saúde no país.¹⁰ A lista dos indicadores disponíveis no TabNet, na página da SES/SP, está reproduzida na Tabela 1. O cálculo dos indicadores segue a metodologia adotada pela Rede Interagencial de Informações para a Saúde (RIPSA) (<http://tabnet.datasus.gov.br/tabdata/livroidb/2ed/indicadores.pdf>).¹¹

Tabela 1 – Lista de indicadores disponíveis no endereço-padrão (<http://portal.saude.sp.gov.br/links/matriz>) do pacote rtabnetsp

Número identificador (ID) no pacote	Nome do indicador	Número de subindicadores	Período disponível
1	1b – População segundo a Fundação Sistema Estadual de Análise de Dados de São Paulo (SEADE)	1	2000-2020
2	2 – Densidade demográfica segundo o Censo Demográfico da Fundação Instituto Brasileiro de Geografia e Estatística (IBGE)	3	2000-2010
3	3 – Grau de urbanização segundo o Censo Demográfico da Fundação Instituto Brasileiro de Geografia e Estatística (IBGE)	3	2000-2010
4	4 – Taxa de crescimento populacional (2000-2010) segundo o Censo Demográfico da Fundação Instituto Brasileiro de Geografia e Estatística (IBGE)	3	2010-2010
5	5 – Proporção de idosos	3	2000-2018
6	6 – Taxa bruta de natalidade	3	2000-2018
7	7 e 8 – Percentual de municípios com menos de 10 mil e com mais de 100 mil habitantes	2	2010-2018
8	Proporção (%) de municípios nos grupos 4 e 5, segundo Região de Saúde, DRS e RRAS do estado de São Paulo – 2008 a 2014 – versão 2016	1	2008-2014
9	10 – Produto interno bruto (PIB) <i>per capita</i> em reais (R\$) corrente	3	2000-2017
10	11 a 13 – Taxa de mortalidade infantil e componentes	11	2000-2019
11	14 – Razão de mortalidade materna	3	2000-2018
12	15 a 17 – Taxa de mortalidade por neoplasias	47	2000-2019
13	18 – Taxa de mortalidade por doenças do aparelho circulatório	6	2000-2019
14	19 – Taxa de mortalidade por causas externas	17	2000-2019
15	20 a 23 – Percentual de óbitos por causas definidas	18	2000-2019
16	24 e 25 – Percentual de nascidos vivos com baixo peso ao nascer	5	2000-2019
17	26 – Percentual de partos em menores de 20 anos de idade	3	2000-2018
18	27a – Taxa de letalidade por formas graves de dengue (FHD + SCD + DCC)	3	2000-2011
19	27b – Taxa de letalidade de dengue	3	2012-2018
20	28 – Taxa de incidência de aids	3	2000-2018
21	29 – Taxa de incidência de sífilis congênita	3	2007-2018
22	30 – Prevalência de pacientes em diálise	3	2000-2017
23	31a – Leitos-SUS por 1.000 (mil) habitantes	3	2005-2019
24	31b – Leitos-SUS por 1.000 (mil) habitantes na população SUS-dependente	3	2005-2017
25	32 – Percentual de leitos-SUS em unidades de terapia intensiva (UTI)	3	2014-2019
26	33 – Cobertura da Atenção Básica em Saúde	3	2006-2018
27	34 – Cobertura de agente comunitário de saúde (ACS)	3	2000-2017
28	35 – Proporção de nascidos vivos de mães com 7 ou mais consultas de pré-natal	3	2000-2019
29	36a – Percentual de internação SUS na população total residente	3	2000-2018
30	36b – Percentual de internação SUS na população total residente para a população SUS-dependente	3	2000-2018
31	37a – Razão de exames citopatológicos do colo do útero em mulheres de 25 a 64 anos de idade	4	2008-2019
32	37b – Razão de exames citopatológicos do colo do útero em mulheres de 25 a 64 anos de idade na população SUS-dependente	4	2008-2019

Continua

Continuação

Tabela 1 – Lista de indicadores disponíveis no endereço-padrão (<http://portal.saude.sp.gov.br/links/matriz>) do pacote rtabnet

Número identificador (ID) no pacote	Nome do indicador	Número de subindicadores	Período disponível
33	38 – Cobertura vacinal tetravalente (até 2012) e pentavalente (a partir de 2013) em menores de 1 ano de idade	3	2000-2018
34	39 – Média de consultas médicas por habitante nas especialidades básicas	3	2000-2017
35	40 – Proporção de consultas de urgência por consulta básica	3	2000-2017
36	41 – Cobertura de 1ª consulta odontológica programática	3	2000-2016
37	42 – Percentual de internação por condições sensíveis à Atenção Básica em saúde (ICSAB)	3	2000-2018
38	43 – Taxa de internação por acidente vascular cerebral (AVC) em maiores de 40 anos de idade	3	2000-2019
39	44 – Taxa de internação hospitalar por fratura de fêmur em maiores de 60 anos de idade	3	2000-2018
40	45a – Percentual de partos cesáreos – Total (Sistema de Informações sobre Nascidos Vivos [Sinasc])	3	2000-2018
41	45b – Percentual de partos cesáreos pelo SUS (Sistema de Informações Hospitalares [SIH/SUS])	4	2000-2018
42	46 – Cobertura de Centro de Atenção Psicossocial (CAPS)	9	2012-2018
43	47a – Cobertura da Saúde Suplementar	3	2000-2019
44	47b – Estimativa da população SUS-dependente (com base na Saúde Suplementar)	4	2000-2019
45	48 – Proporção de cura de casos novos de tuberculose pulmonar bacilífera	3	2000-2018
46	49 – Proporção de cura de casos novos de hanseníase diagnosticados	3	2001-2017
47	50 a 52 – Índice de qualidade do tratamento da água nos parâmetros: bacteriológico, cloro e flúor	9	2008-2016
48	53 – Despesa em saúde <i>per capita</i>	3	2000-2017
49	54a – Razão de mamografia de rastreamento em mulheres de 50 a 69 anos de idade	4	2010-2018

Pacote	Hiperlink	Função
httr	https://cran.r-project.org/web/packages/httr/httr.pdf	Realização das requisições HTTP ao servidor do TabNet
xml2	https://cran.r-project.org/web/packages/xml2/xml2.pdf	Scraping das páginas
rvest	https://cran.r-project.org/web/packages/rvest/rvest.pdf	
stringi	https://cran.r-project.org/web/packages/stringi/stringi.pdf	Manipulação de texto e tratamento de codificação para o padrão Windows-1252 (https://encoding.spec.whatwg.org/#windows-1252)
purrr	https://cran.r-project.org/web/packages/purrr/purrr.pdf	Tratamento de exceções nas requisições HTTP
tidyr	https://cran.r-project.org/web/packages/tidyr/tidyr.pdf	Manipulação e tratamento de dados
dplyr	https://cran.r-project.org/web/packages/dplyr/dplyr.pdf	
ggplot2	https://cran.r-project.org/web/packages/ggplot2/ggplot2.pdf	Construção de gráficos coropléticos simples, a partir dos dados obtidos
sf	https://cran.r-project.org/web/packages/sf/sf.pdf	Representação cartográfica
RColorBrewer	https://cran.r-project.org/web/packages/RColorBrewer/RColorBrewer.pdf	Paleta de cores, para representação cartográfica

Figura 1 – Dependência de outros pacotes R utilizados no rtabnet

No `rtabnet`, foram reunidas funções práticas para visualização, exploração e obtenção dos dados de indicadores de saúde do TabNet da SES/SP. O pacote está disponível na plataforma do GitHub (<https://github.com/>), pelo *link* <https://github.com/joaohmorais/rtabnet>

Quanto aos aspectos técnicos, o `rtabnet` utiliza dependências de outros pacotes, listados na Figura 1.

Instalação do software R

Para instalar o R, deve-se acessar sua página eletrônica (<https://www.r-project.org/>) e seguir as instruções de instalação. Antes de passar à próxima etapa, requer-se a instalação do pacote `devtools` (<https://cran.r-project.org/web/packages/devtools/devtools.pdf>) no R.

Instalação do pacote `rtabnet` no R

O pacote `rtabnet` deve ser instalado por meio de seu repositório no *website* GitHub (<https://github.com/joaohmorais/rtabnet>). Com o R aberto, digita-se:

```
devtools::install_github("joaohmorais/rtabnet")
```

Para carregar o pacote `rtabnet` no R

Para começar a utilizar o `rtabnet`, carrega-se o pacote digitando-se:

```
library("rtabnet")
```

Isto faz com que o pacote `rtabnet` esteja disponível para uso. Esse passo precisa ser realizado após cada inicialização do R.

Funções do pacote `rtabnet`

O `rtabnet` obtém os indicadores de saúde, por meio da *webpage* da SES/SP. Os indicadores de saúde são calculados mediante fórmulas padronizadas pela RIPSAs,⁷ e podem ser carregados segundo município, Departamento Regional de Saúde (DRS), Região de Saúde (RS) e Redes Regionais de Atenção à Saúde (RRAS) do estado de São Paulo.

A listagem completa dos indicadores disponibilizados pelo `rtabnet` é obtida digitando-se a linha de comando:

```
tabnet_index()$Nomes
```

Função `indicator_list()`

```
indicator_list(url= "http://portal.saude.sp.gov.br/links/matriz")
```

A função `indicator_list()` recupera a lista de indicadores disponíveis na página do TabNet, com seus respectivos números identificadores, denominados IDs, utilizados para escolha dos indicadores. Essa função, quando executada, explora uma determinada URL – definida, por padrão, como o endereço do TabNet da SES/SP – e recupera nomes e *links* de indicadores de saúde identificados na página.

Função `indicator_search()`

```
indicator_search(keywords, url = "http://portal.saude.sp.gov.br/links/matriz")
```

Pode-se também realizar consultas à lista de indicadores de uma página, quando se busca por um indicador específico. A função `indicator_search()` mostrará uma lista com o nome e o ID de cada indicador, e o título correspondente ao termo informado no parâmetro “`keywords`”. Cumpre alertar que a função não distingue letras maiúsculas de minúsculas, mas, sim, diferencia caracteres acentuados de não acentuados.

Função `view_indicador()`

```
view_indicador(indicator_index, url = "http://portal.saude.sp.gov.br/links/matriz", timeout = 1)
```

Uma vez definido o indicador que se deseja visualizar, é importante verificar seus níveis de agregação, períodos e conteúdos disponíveis, antes de realizar uma requisição por dados. A função `view_indicador()` retorna uma lista com informações de níveis de agregação, períodos (anos) e conteúdo. A propósito, um indicador pode ter diferentes conteúdos, diferenciados por subindicadores: por exemplo, o indicador ‘taxa de mortalidade por causas externas’ tem subindicadores, como ‘óbitos por homicídio’, ‘óbitos por suicídio’ e outras causas, disponíveis no TabNet. Para utilizar tal função, deve-se informar como “`indicator_index`” o número de ID do indicador.

O parâmetro “`timeout`” aparece em diversas funções e representa o tempo, em segundos, durante o qual a conexão permanecerá aberta para obtenção de informações da página do indicador, de forma a evitar

a insistência em conexões sem retorno. O “timeout” é padronizado em 1 segundo, embora esse tempo deva se ajustar em caso de conexões lentas.

Função `indicator_df()`

```
indicator_df(indicator_index, region = NULL,
             subindicador = NULL, years = NULL, onlyMostRecent
             = FALSE, url = "http://portal.saude.sp.gov.br/links/
             matriz", timeout = 1)
```

A função `indicator_df()` é a que, de fato, realiza a requisição por dados de um indicador em determinado período e nível de agregação. Uma vez que se tem o ID do indicador, ele é informado à função como “`indicator_index`”. O próximo passo será informar o nível de agregação no parâmetro “`region`”, utilizando-se o termo ‘Município’, ‘Região de Saúde’, ‘RRAS’ ou ‘DRS’. Se isso não for feito, automaticamente a função tomará, como padrão, o primeiro nível disponível (‘Município’, na maioria dos casos). De forma similar, pode-se especificar o conteúdo do indicador (subindicador) no parâmetro “`subindicator`”, e o período desejado, em “`years`”; caso não se especifique esse conteúdo, a função considerará o último subindicador disponível para o indicador, e para todos os períodos disponíveis. A função ainda conta com o parâmetro lógico (os valores

“TRUE” ou “FALSE”) “`onlyMostRecent`”, o qual, se verdadeiro (“TRUE”), recuperará apenas as observações do mais recente entre os anos disponíveis, ou entre os especificados pelo parâmetro “`years`”.

Executada a função, o usuário receberá uma matriz de dados com o código identificador e o nome de cada região, o ano de observação e o valor do indicador observado.

Função `fetch_all()`

```
fetch_all(region = “Município”, url = “http://portal.
saude.sp.gov.br/links/matriz”, timeout = 1)
```

Em alguns tipos de análises, deseja-se trabalhar com vários ou mesmo todos os indicadores disponíveis. A função `fetch_all()` supre essa necessidade, ao agrupar todos os indicadores disponíveis para uma mesma matriz de dados. É possível especificar novamente, através do parâmetro “`region`”, um nível de regionalização desejado. Como resposta, o usuário obtém uma nova matriz de dados com o número identificador e o nome de cada região, o ano de observação, os nomes do indicador e do subindicador em questão, e o valor observado. É importante notar que a função atende a diversas requisições do servidor, uma para cada indicador, e pode levar alguns minutos de execução.

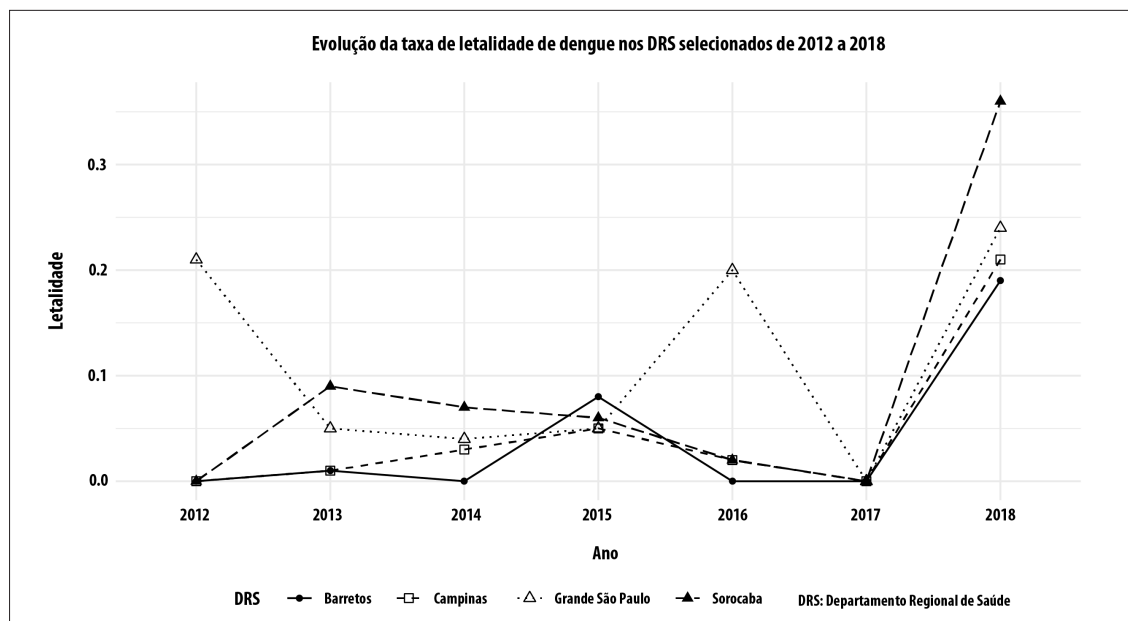


Figura 2 – Exemplo de gráfico de linhas, construído com dados de letalidade de dengue obtidos pelo `rtabnetsp`, referentes a quatro DRS de São Paulo, 2012-2018

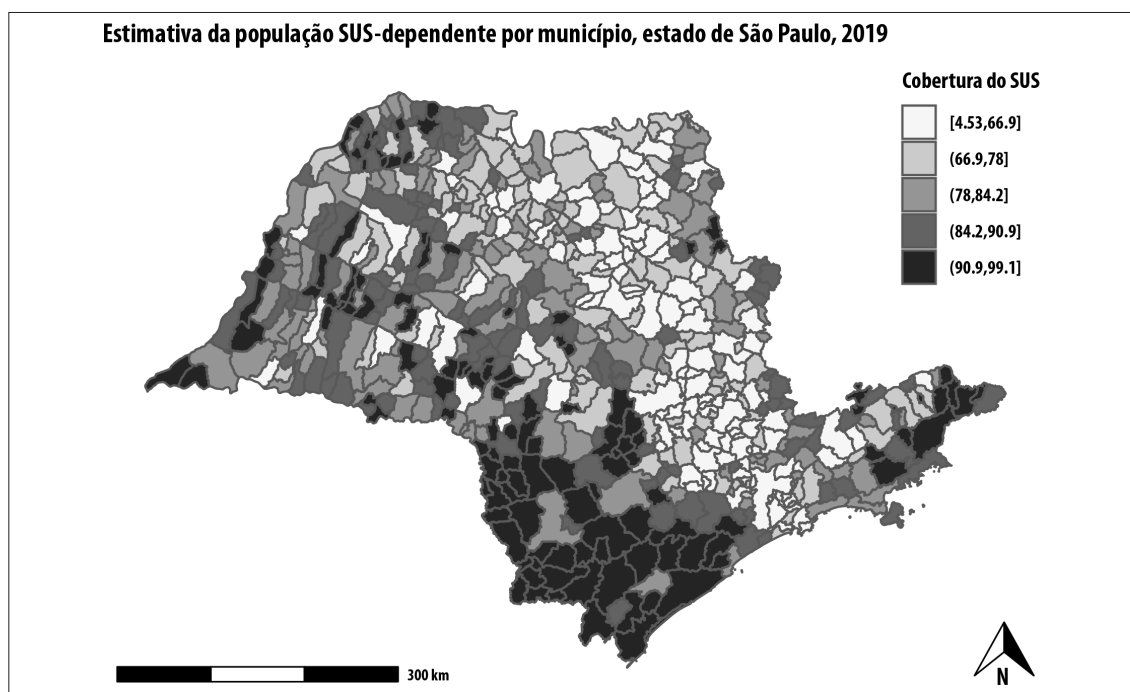


Figura 3 – Mapa coroplético da população dependente do Sistema Único de Saúde (SUS), por município de São Paulo, desenhada pelo pacote rtabnet

Função tabnet_map()

```
tabnet_map(indicator_index, region = "Município",
           subindicator = NULL, years = NULL, label = FALSE,
           nBins = NULL, palette = "Purples", title = NULL)
```

Por último, existe a possibilidade de obter-se um mapa coroplético, a partir dos dados de algum indicador, pela função `tabnet_map()`, cujos parâmetros são similares aos da função `tabnet_df()`, em que se pode especificar o ID do indicador, o nível de agregação, o conteúdo e o período desejados. Adicionalmente, essa função apoia os seguintes parâmetros: “label”, que pode ser verdadeiro ou falso, indicando se o usuário deseja os nomes das regiões no mapa; “nBins”, pelo qual é possível especificar o número de intervalos (entre 3 e 9) em que serão divididos os valores do indicador; “palette”, que permite a personalização da paleta de cores do mapa, de acordo com o pacote `RColorBrewer` (<https://cran.r-project.org/web/packages/RColorBrewer/RColorBrewer.pdf>); e “title”, para a definição do título do mapa.

Considerações éticas

O pacote `rtabnet` foi desenvolvido no âmbito do projeto ‘Desenvolvimento de aplicativo R/Shiny para auxílio na visualização e análise de indicadores e análise de indicadores de saúde no estado de São Paulo’, este aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de São Paulo (CEP/Unifesp): Certificado de Apresentação para Apreciação Ética (CAAE) nº 94098718.7.0000.5505, emitido em 20 de setembro de 2018.

Resultados

Nesta seção, apresentam-se situações hipotéticas comuns ao contexto de gestão em saúde, exemplos ilustrativos das funções do `rtabnet` utilizadas no processo.

Situação 1 – Visualização temporal da taxa de letalidade de dengue por DRS

Neste primeiro cenário, deseja-se visualizar a evolução da taxa de letalidade de dengue desde 2012,

em quatro DRS específicos: Grande São Paulo, Campinas, Barretos e Sorocaba. Com o `rtabnet` carregado no R, o primeiro passo será encontrar o indicador desejado:

```
> indicator_search("dengue")
[1] "18 – Taxa de letalidade por formas graves de dengue até 2011 (FHD+SCD+DCC)"
[2] "19 – Taxa de letalidade de dengue (a partir de 2012)"
```

Como se trata de histórico recente, opta-se pelo indicador ID 19 – Taxa de letalidade de dengue (a partir de 2012). Em seguida, verificam-se os níveis de agregação e períodos disponíveis para o indicador:

```
> view_indicator(19)
$Indicator.Name
[1] "19 – Taxa de letalidade de dengue"
$Indicator.URL
[1] "http://tabnet.saude.sp.gov.br/defthtm.exe?tabnet/ind27b_matriz.def"
$Available.Regions
[1] "Município" "DRS" "DRS/Reg Saude" "DRS/Mun" "RRAS" "RRAS-Reg Saude" "RRAS-Município"
[8] "Região de Saúde" "Reg Saúde/Mun"
$Available.Years
[1] "2018" "2017" "2016" "2015" "2014" "2013" "2012"
$Available.Subindicators
[1] "Óbitos" "Total de Casos" "Tx Letalidade Dengue"
```

A partir destas informações, verifica-se a disponibilidade de dados do indicador para todos os anos de 2012 a 2018, assim como a regionalização desejada (DRS). O subindicador desejado é o último da lista: a taxa de letalidade. Logo, utiliza-se a função `indicator_df()` para recuperar os dados do indicador e salvá-los em uma matriz de dados:

```
> dados <- indicator_df(19, region = "DRS")
> head(dados)
id DRS Ano Valor
1 3501 Grande São Paulo 2012 0.21
2 3502 Araçatuba 2012 0.07
3 3503 Araraquara 2012 0.00
4 3504 Baixada Santista 2012 0.09
5 3505 Barretos 2012 0.00
6 3506 Bauru 2012 0.00
```

Assim, tem-se uma matriz de dados salva no elemento "dados", com os códigos identificadores e nomes dos DRS, o ano de ocorrência e o valor do indicador naquele ano e naquele DRS. Para visualizar a tendência do indicador ao longo dos anos, entre os DRS desejados, pode-se utilizar um gráfico de linhas. O pacote `ggplot2` permite esse recurso, a partir da matriz de dados obtida. O resultado é ilustrado na Figura 2: visualiza-se a evolução das taxas de letalidade de dengue no período de 2012 a 2018, nos DRS selecionados no recorte: Grande São Paulo, Campinas, Barretos e Sorocaba. Nota-se, por exemplo, uma alta na taxa de letalidade de dengue na Grande São Paulo, em 2016, e em Sorocaba, no ano de 2018.

```
recorte.DRS <- c("Grande São Paulo", "Campinas", "Barretos", "Sorocaba")
dados <- dados[dados$DRS %in% recorte.DRS,]
ggplot(data=dados, aes(x=Ano, y=Valor, group=DRS)) +
  geom_line(aes(linetype = DRS)) +
  geom_point(aes(shape = DRS)) +
  theme_minimal() +
  theme(legend.position = "bottom", plot.title = element_text(hjust = 0.5)) +
  labs(x = "Ano", y = "Letalidade") +
  ggtitle("Evolução da taxa de letalidade de dengue nos DRS selecionados, \nde 2012 a 2018")
```

Situação 2 – Visualização espacial da população SUS-dependente

Noutro cenário, deseja-se visualizar a distribuição espacial da cobertura da população SUS-dependente entre os municípios do estado de São Paulo. De forma similar à situação hipotética anterior, deve-se localizar o número identificador do indicador desejado:

```
> indicator_search("dependente")
[1] "24 – Leitos SUS por 1.000 habitantes para a população SUS-dependente"
[2] "30 – Percentual de internação da população residente para a população SUS-dependente"
[3] "32 – Razão de exames citopatológicos cervico-vaginais para a população SUS-dependente"
[4] "44 – Estimativa da população SUS-dependente (com Base na Saúde Suplementar)"
```

Pode-se utilizar a função `tabnet_map()` para o desenho de um mapa coroplético do indicador.

Por padrão, a função recuperará dados no nível de município, e no ano mais recente disponível. Logo, não há necessidade de especificar, por meio dos parâmetros da função, nem o nível de agregação, nem o período desejado. Reitera-se a necessidade da dependência do pacote `sf` para essa função. A Figura 3 mostra o mapa gerado pela função, apresentando a distribuição espacial da cobertura da população SUS-dependente.

```
> tabnet_map(44, nBins = 5, palette = "Greys", title = "Estimativa da população SUS dependente por município. Estado de São Paulo, 2019.")
```

Discussão

O pacote `rtabnet` torna prática a recuperação de dados de indicadores de saúde disponíveis na página da Secretaria de Estado da Saúde de São Paulo, diretamente para o *software* R, pelo qual podem ser manipulados. Mediante ambas as situações apresentadas, a título de exemplo, foi possível acompanhar o emprego de recursos de análise descritiva com poucos comandos, dentro do programa. O desenvolvimento do pacote `rtabnet` atende ao propósito da vigilância em saúde de aumentar a acessibilidade à informação, possibilitando análises sistemáticas na área de gestão em saúde.

O `rtabnet` difere de outros pacotes já abordados na literatura, quando se trata de recuperação de informações dos sistemas do Ministério da Saúde. O pacote `microdatasus`,⁵ por exemplo, obtém dados de

ocorrência a partir dos arquivos DBF, disponíveis em cada base disponibilizada pelo Datasus.

O pacote objeto deste estudo desenvolve uma operação de *web scraping* para recuperação de dados já tabulados pela ferramenta TabNet existente. O método é mais vantajoso em análises de dados agregados, imprimindo mais agilidade e confiabilidade ao processo, uma vez que os dados já se encontram tabulados pelo sistema do TabNet.

Acredita-se que o desenvolvimento do `rtabnet` cumpre seu objetivo principal: facilitar o acesso e a análise de indicadores e auxiliar a tomada de decisão em saúde, no estado de São Paulo. O pacote ainda reúne o potencial de contribuir com o desenvolvimento de outros *softwares* projetados para a abordagem de dados de saúde, e a possibilidade de sua utilização como mediador para aplicativos que necessitam de tais dados e busquem o acompanhamento de indicadores.

Contribuição dos autores

Morais JHA e Martins CB contribuíram na concepção e delineamento do estudo, análise e interpretação dos resultados, redação e revisão crítica do conteúdo do manuscrito. Konstantyner TCRO contribuiu na análise e interpretação dos dados, redação e revisão crítica do conteúdo do manuscrito. Fazenda AL e Sala A contribuíram na concepção e delineamento do estudo, redação e revisão crítica do conteúdo do manuscrito. Todos os autores aprovaram a versão final do manuscrito e são responsáveis por todos os seus aspectos, incluindo a garantia de sua precisão e integridade.

Referências

1. Hartz ZMA, Silva LMV. Avaliação em saúde: dos modelos teóricos à prática na avaliação de programas e sistemas de saúde [Internet]. Salvador: EDUFBA; Rio de Janeiro; Editora FIOCRUZ; 2005 [citado 2020 jun 19]. Disponível em: <http://books.scielo.org/id/xzdnf>
2. van Panhuis WG, Paul P, Emerson C, Grefenstette J, Wilder R, Herbst AJ, et al. A systematic review of barriers to data sharing in public health. BMC Public Health [Internet]. 2014 Dec [cited 2020 Aug 22];14(1). Available from: <https://doi.org/10.1186/1471-2458-14-1144>
3. Health Metrics Organization. World Health Organization - WHO. Framework and standards for country health information systems [Internet]. 2nd ed. Geneva: World Health Organization; 2008 [cited 2020 Aug 23]. Available from: <https://apps.who.int/iris/handle/10665/43872>
4. Ministério da Saúde (BR). Datasus: trajetória 1991 – 2002 [Internet]. Brasília: Ministério da Saúde; 2002 [citado 2020 jun 17]. Disponível em: http://www.saude.gov.br/bvs/publicacoes/trajetoria_datasus.pdf
5. Saldanha RF, Bastos RR, Barcellos C. Microdatasus: pacote para download e pré-processamento de microdados do Departamento de Informática do SUS (DATASUS). Cad Saúde Pública [Internet]. 2019 set [citado 2020 jun 17];35(9). Disponível em: <https://doi.org/10.1590/0102-311x00032419>

6. R Development Core Team. A language and environment for statistical computing: reference index [Internet]. Vienna: R Foundation for Statistical Computing; 2010 [cited 2020 Jun 17]. Available from: <http://www.polsci.wvu.edu/duval/PS603/Notes/R/fullrefman.pdf>
7. Nedel FB. Pacote csapAIH: a lista brasileira de internações por condições sensíveis à atenção primária no programa R. *Epidemiol Serv Saúde* [Internet]. 2019 [citado 2020 ago 24];28(2):e2019084. Disponível em: <https://doi.org/10.5123/s1679-49742019000200021>
8. Rocha TAH, Silva NC, Amaral PVM, Barbosa ACQ, Vissoci JRN, Thomaz EBAF, et al. Geolocalização de internações cadastradas no Sistema de Informações Hospitalares do Sistema Único de Saúde: uma solução baseada no programa estatístico R. *Epidemiol Serv Saúde* [Internet]. 2018 [citado 2020 ago 24];27(4):e2017444. Disponível em: <https://doi.org/10.5123/s1679-49742018000400016>
9. Viacava F, Ugá MAD, Porto S, Laguardia J, Moreira RS. Avaliação de desempenho de sistemas de saúde: um modelo de análise. *Ciênc Saúde Coletiva* [Internet]. 2012 abr [citado 2020 nov 26];17(4):921-34. Disponível em: <https://doi.org/10.1590/S1413-81232012000400014>
10. Barcellos CC, Sabroza PC, Peiter P, Iñiguez Rojas L. Organização espacial, saúde e qualidade de vida: análise espacial e uso de indicadores na avaliação de situações de saúde. *Inf Epidemiol SUS* [Internet]. 2002 set [citado 2020 jun 17];11(3):129-38. Disponível em: <https://www.arca.fiocruz.br/handle/icict/713>
11. Rede Interagencial de Informações para a Saúde, organizador. Indicadores básicos para a saúde no Brasil: conceitos e aplicações [Internet]. 2. ed. Brasília: Organização Pan-Americana da Saúde; 2008 [citado 2020 nov 26]. 349 p. Disponível em: <http://tabnet.datasus.gov.br/tabdata/livroidb/2ed/indicadores.pdf>

Abstract

Health status indicators are an important tool for monitoring the performance of public health actions, identifying trends and priority regions for resource allocation. An R package was developed in order to increase the feasibility of handling and analyzing health status indicator data. The rtabnet package requests data from TabNet servers on the São Paulo State Department of Health website, retrieving and preprocessing the data for user manipulation. This article presents the rtabnet package and its functions, installation and use; as well as providing examples of its functionalities, which involve listing and searching among available indicators, selecting desired content and obtaining data aggregated according to regionalization level held on the data matrix, enabling greater agility in tasks regarding public health management in the state of São Paulo.

Keywords: Software; Electronic Data Processing; Health Status Indicators; Health Information Systems; Health Management.

Resumen

Indicadores de salud son una herramienta importante para monitorear el desempeño de las acciones de salud pública, permitiendo la evaluación de las intervenciones hechas, así como la identificación de tendencias y regiones prioritarias para la asignación de recursos. En la búsqueda de aumentar la practicidad en las tareas de análisis y manipulación de datos de estos indicadores, se creó un paquete R. El paquete rtabnet realiza solicitudes a los servidores TabNet del Departamento de Salud del Estado de São Paulo, recogiendo y procesando dichos datos para el usuario. Este artículo presenta el paquete, sus funciones, instalación y uso, así como ejemplos de sus funcionalidades, que permiten visualizar y buscar desde un listado de indicadores, seleccionar el contenido deseado y obtener los datos agregados por el nivel de regionalización disponible en la matriz de datos, alcanzando más agilidad en las tareas de gestión de la salud en el estado de São Paulo.

Palabras clave: Programas Informáticos; Procesamiento Automatizado de Datos; Indicadores de Salud; Sistemas de Información en Salud; Gestión en Salud.

Recebido em 24/08/2020
Aprovado em 20/10/2020

Editor associado: Lúcia Rolim Santana de Freitas – orcid.org/0000-0003-0080-2858