

# Como desenvolve-R um Índice R-reprodutível e Compa-R-ável

**Autor(as):** Aline da Nóbrega Oliveira, Rogerio Vidal de Siqueira, Leandro de Almeida Salles, Rafael de Acypreste Monteiro Rocha, Luiz Antonio Gouveia de Oliveira

**Palabras clave:** Desempenho Ambiental Urbano, Sustentabilidade, Políticas Públicas, Reprodutibilidade

## Abstract

O desenvolvimento do Índice de Sustentabilidade Urbana do Distrito Federal envolveu a formulação de uma metodologia capaz de criar um índice comparável ao longo do tempo para permitir uma análise mais precisa do impacto das políticas públicas urbanas e ambientais. Utilizando a linguagem R, o índice incorpora bases de dados de órgãos e entidades do DF, assim como valores de referência da literatura, assegurando que os indicadores fossem avaliados de forma consistente entre as 35 Regiões Administrativas (RAs) do DF. A flexibilidade e reproduzibilidade do R permitiram testar diferentes cenários e ajustar o índice para que refletisse com precisão a realidade urbana ambiental do DF, garantindo sua continuidade e eficácia na avaliação das políticas públicas ao longo dos ciclos de governo.

## Introdução

O Índice de Sustentabilidade Urbana foi desenhado para ser uma ferramenta robusta e comparável ao longo do tempo, permitindo uma análise mais abrangente do comportamento socioambiental das áreas urbanas das 35 Regiões Administrativas (RAs) do Distrito Federal. O desenvolvimento desse Índice avalia o impacto da sociedade sobre a dinâmica urbano ambiental, bem como o efeito das políticas públicas ao longo de sucessivos períodos governamentais. proporcionando uma ferramenta crucial para o monitoramento contínuo da sustentabilidade ambiental

## Metodologia

A implementação do Índice de Sustentabilidade Urbana foi facilitada pelo uso da linguagem R, organizando as diferentes bases provenientes de órgãos e entidades do Distrito Federal, etapas de processamento, cálculo e visualização dos dados com pacotes específicos. A importação e exportação dos dados foi realizada diretamente de arquivos

Excel utilizando o pacote *openxlsx*. Após a importação, os dados foram limpos e organizados usando as funções do *tidyverse*. Com os dados organizados, o cálculo dos indicadores foi realizado utilizando tanto funções do R base quanto do *dplyr*. Em sua formulação, o índice abrange diferentes dimensões temáticas - áreas verdes e uso do solo, saneamento ambiental, infraestrutura e mudanças climáticas - essas dimensões foram calculadas individualmente para cada Região Administrativa (RA), utilizando a função *group\_by()* do *dplyr*, que permitiu a agregação dos indicadores por RA. Para apresentação dos resultados foram utilizados os pacotes *gt* para tabelas e *ggplot2* para gráficos, o que permitiu uma abordagem dinâmica e flexível na estruturação dos indicadores. A estruturação dos dados e cálculos em R garantiu que o índice pudesse ser facilmente atualizado e reproduzido, à medida que novas informações se tornassem disponíveis. Por fim, o script foi configurado para importar automaticamente novas bases de dados diretamente da rede, utilizando a função *download.file()* do R base, o que permitiu que os resultados fossem gerados em tempo real. Isso garantiu que o índice fosse atualizado de forma contínua, à medida que novas informações fossem disponibilizadas pelos órgãos competentes.

## Resultados

A partir do desenvolvimento metodológico facilitado pelo R, as tabelas geradas com *gt()* mostraram comparações precisas dos indicadores entre as regiões, enquanto os gráficos criados com *ggplot()* destacaram variações ambientais e socioeconômicas significativas, com algumas RAs apresentando altos níveis de desempenho em determinadas dimensões, enquanto em outras áreas mostravam desafios socioambientais mais acentuados. O uso do *tidyverse* no tratamento dos dados e na exploração de diferentes cenários foi fundamental para testar diferentes pesos e estruturas de cálculo. A função *mutate()* foi amplamente utilizada para ajustar os indicadores, o que garantiu a flexibilidade na formulação do índice. Isso permitiu a avaliação de cenários com variação de pesos entre dimensões, assegurando que o índice fosse representativo das realidades socioambientais de cada RA. Isso assegura que o índice reflita de maneira mais precisa as condições e desafios da sustentabilidade local, tornando-se uma ferramenta essencial para o monitoramento contínuo e a adaptação das políticas públicas.

## Conclusão

O uso da linguagem R na elaboração do Índice de Sustentabilidade representa um avanço significativo para o direcionamento das ações de Governo e suas políticas públicas. A metodologia baseada em R garante que o índice seja uma ferramenta eficaz, reprodutível e facilmente atualizável, adequando-se às transições de equipes entre governos. Além disso, a metodologia que pode ser alimentada por bases de dados diversas e a reprodutibilidade e gratuidade proporcionadas pelo R facilitam o acesso de instituições de outras unidades da federação a esta ferramenta, que tem se mostrado essencial para avaliar a sustentabilidade

## **COMO DESENVOLVE-R UM ÍNDICE R-EPRODUTÍVEL E COMPA-R-ÁVEL**

---

territorial, tema de grande relevância nas agendas globais da sustentabilidade em prol do enfrentamento das mudanças climáticas nas cidades.