

DASHBOARD PARA TOMADA DE DECISÃO SOBRE SINISTROS DE TRÂNSITO USANDO FRAMEWORK SHINY VIA LINGUAGEM R.

Mário Diego Rocha Valente ¹

Kleber Bezerra Salim²

Tito Félix Oliveira³

Breno Cauã Rodrigues da Silva⁴

Resumo

O Departamento de Trânsito do Estado do Pará (DETRAN-PA) é responsável pela consolidação, análise e divulgação dos dados oficiais de sinistros de trânsito do Estado do Pará, com base em dados fornecidos pelos Boletins de Ocorrência da Polícia Civil, do Batalhão Rodoviário da Polícia Militar (BOAT), Laudos Periciais (IML), informações de Resgate (SAMU e BOMBEIRO). O Processo de Integração de Múltiplas bases de dados foi realizado utilizando o Método Probabilístico de Relacionamento de Registro chamado de Pareamento (*linkage*) desenvolvido por Fellegi e Sunter, (1969). Com a integração dos dados foi criada uma Lista única de Vítimas Fatais decorrente de Sinistros de Trânsito, utilizando um processo de Extração-Tratamento-Carga. Dessa forma, o R e todo o *framework Shiny* surgem com um enorme potencial de atender aos públicos internos e externos, uma ferramenta open *Source* para substituir as rotinas anteriores feitas em software empresarial (SPSS) com limitações (IBM, 2023). Para a criação do dashboard foi implementado um script na linguagem R em um ambiente de desenvolvimento integrado Rstudio, com uso do pacote *bs4Dash* para disponibilização de dados de segurança viária tanto para órgãos do governo quanto para a população. **Palavras-chave:** Sinistro de Trânsito, DETRAN-PA, Segurança Viária, Shiny, R.

Abstract

The Traffic Department of the State of Pará (DETRAN-PA) is responsible for consolidating, analyzing, and disseminating official traffic accident data in the state. To enhance this process, a multi-database integration model was implemented using the Probabilistic Record Linkage Method developed by Fellegi and Sunter (1969). This integration resulted in the creation of a unique list of fatal traffic accident victims through an Extract-Transform-Load (ETL) process. Previously, these analyses were conducted using enterprise software (SPSS) with various limitations. In this context, the R programming language and the Shiny framework emerge as efficient and open-source alternatives for interactive data visualization. As a result, an interactive platform was developed using Shiny to provide road safety data to

¹Departamento de Trânsito do Estado do Pará (DETRAN-PA), mario.valente@detran.pa.gov.br

²Departamento de Trânsito do Estado do Pará (DETRAN-PA), kleber.salim@detran.pa.gov.br

³Departamento de Trânsito do Estado do Pará (DETRAN-PA), tito.felix@detran.pa.gov.br

⁴Universidade Federal do Pará (UFPA), aluno de graduação, brenosilvasantos25@gmail.com

both government agencies and the general public. The tool enables dynamic access and intuitive data visualization, promoting greater transparency and supporting evidence-based public policies aimed at reducing traffic accidents. **Keywords:** Traffic Accident, DETRAN-PA, Road Safety, Shiny, R.

Introdução

O Departamento de Trânsito do Estado do Pará (DETRAN-PA) é uma autarquia responsável pela consolidação, análise e divulgação dos dados oficiais de segurança viária do estado do Pará (DETRAN, 2024a). Diariamente, são extraídos os dados dos Boletins de Ocorrência da Polícia Civil e Boletins do Batalhão Rodoviários de Polícia Militar, estes são consolidados por uma equipe multidisciplinar que inclui estatísticos, economistas e técnicos de segurança pública, e até então divulgados por meio de bases de dados temáticas e planilhas para a imprensa.

Durante os últimos 5 anos, uma das atualizações realizadas pela gerência de estatística de trânsito, vinculada à coordenadoria de planejamento, foi a substituição de softwares pagos como o caso SPSS (IBM CORP, 2020) para soluções em linguagem de programação R (R Core Team, 2024).

A priori, estas soluções buscavam simplificar rotinas e automatizar processos, porém, com o tempo, foram gradativamente expandidas para a produção de relatórios em markdown, até chegarem a construção de painéis de dados (dashboard) através de um servidor Ubuntu configurado no GitLab.

A criação dos dashboards desenvolvidos no pacote *Shiny* (Chang *et al.*, 2024) tem como objetivo principal a substituição da antiga infraestrutura de painéis hospedados no Excel. Por se tratar de uma opção open source, o *Shiny* apresenta inúmeras opções de construção de interfaces e layout para o usuário, possibilidades de automação e conexão com diferentes tipos de bases de dados e integração com todo o ambiente de pacotes de análise de dados do R.

O dashboard apresenta informações já trabalhadas e mais estruturadas de que os bancos de dados disponíveis atualmente pelos órgãos públicos, a fim de reduzir os transtornos causados pelos sinistros de trânsito, como por exemplo decidir sobre determinado trecho a ser mais fiscalizado por Agentes de Trânsito durante um período de tempo.

De acordo com a problemática, este artigo visa apresentar uma ferramenta (em formato de dashboard) que transforma microdados referentes a sinistros de trânsito disponibilizados pelo DETRAN-PA em informação disponível à população em geral.

Objetivo

Relatar a experiência do desenvolvimento de um dashboard para Sinistros de Trânsito, uma plataforma baseada em Linguagem de Programação R, usando o pacote Shiny e

sistema de banco de dados relacional **ORACLE** para hospedar a nova infraestrutura de dashboards do Departamento de Trânsito do Estado do Pará, de modo a potencializar a capacidade da Autarquia de Trânsito em publicizar os dados, com maior eficiência e menor custo operacional.

Material e Método

As principais fontes de dados do Departamento de Trânsito do Estado do Pará, são as informações dos boletins de ocorrências registrados nas delegacias de Polícia Civil e extraídos diariamente do Sistema Integrado de Segurança Pública (SISP). Este dado contém informações administrativas, contendo o horário, dia, mês e local em que o registro foi realizado, a tipificação e gravidade do sinistro de trânsito, informações sobre a dinâmica de ocorrência dos eventos, tais como, a condição e natureza da vítima, os tipos de veículos envolvidos, as principais causas contributivas para o desfecho.

Todos os gráficos (pizza, barra, linha) apresentados nos painéis são construídos com o pacote **echarts4r** (Cone, 2023), que possibilita a construir gráficos interativos utilizando uma biblioteca em *Java script*. A utilização do **echarts4r** como pacote para construir os gráficos trouxe diversas opções de interatividade, usando a função **Mouseover** para revelar outras informações ao passar o mouse sobre a figura, com possibilidade de fazer download dos gráficos e Mapas.

Em seguida é apresentado a base de dados em formato de tabela dinâmica, criado com o pacote **DT** (XIE et al, 2025), que fornece uma interface R para a biblioteca *Java Script Datatables*. A função principal neste pacote é *datatables()*, criando um widget HTML para exibir objetos de dados no R, apresentando diversas opções de interatividade, e permitindo a exportação das tabelas em diferentes formatos como CSV, XLS ou PDF.

Em seguida, a aba de distribuição espacial apresenta os Mapas Coropléticos interativos construídos via pacote **Leaflet** (Chang et al., 2024) para diferentes divisões espaciais do estado, no caso, por municípios e regiões de integração.

O código completo do aplicativo pode ser encontrado disponível online no repositório do GitHub, plataforma de hospedagem e compartilhamento de códigos e projetos de programação online (GitHub, 2024).

Resultados e Discussão

Para desenvolver um dashboard sobre sinistros de trânsito, é necessário primeiramente, realizar o processo de integração de múltiplas bases de dados, haja vista que os dados provêm de diversas fontes no Estado do Pará, tais como: Resgate, Saúde, Policiamento e Trânsito. O processo foi realizado seguindo 4 etapas.

A primeira etapa é o diagnóstico situacional, identificando todas as instituições que possuem dados sobre sinistros de trânsito no Estado do Pará. Durante esta etapa, as

informações observadas são: nome da organização; se atende a cena da ocorrência; se gera boletim físico ou eletrônico, se tem acesso ao banco de dados e a natureza dos dados (da cena do sinistro, hospitalar ou mortalidade).

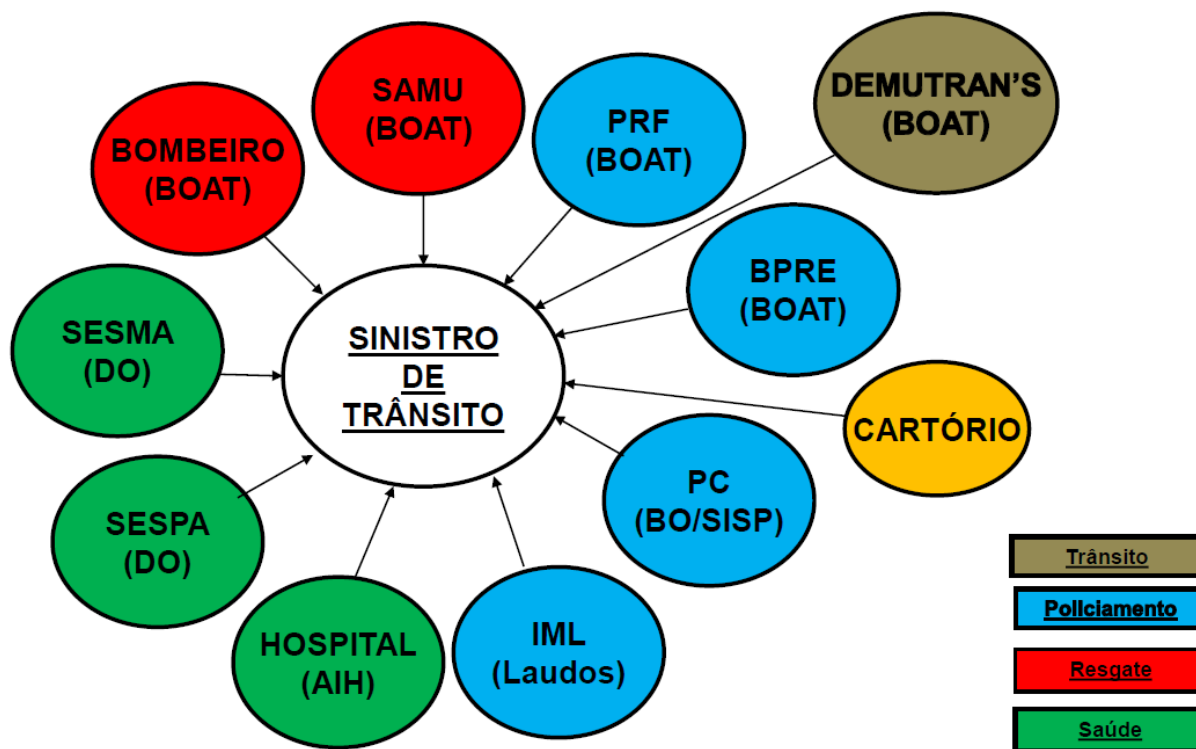


Figura 1: Tipificação de Fontes de Dados sobre Sinistros de Trânsito identificadas e mapeados para o Estado do Pará, versão março de 2025.

Fonte: DETRAN-PA, 2024

Em seguida são definidas quais características (variáveis) de cada base de dados deverão ser compartilhadas, a fim de se construir uma base integrada de vítimas de sinistros de trânsito. Com isso, serão utilizados os dados como nome, sexo e idade/data de nascimento e condição da vítima, endereço, data e horário da ocorrência.

O segundo passo é a **Limpeza, Padronização**, na qual são identificadas as variáveis comuns entre as bases de dados. Tal procedimento visa viabilizar a remoção de registros duplicados relacionados a mesma vítima de uma determinada ocorrência.

A padronização foi usada para a preparação das variáveis e diminuição de erros para a fase de pareamento. Foram retirados acentos, cedilhas, espaços, Algarismo e outros símbolos, e feita a conversão dos nomes das vítimas de letras minúsculas em maiúsculas.

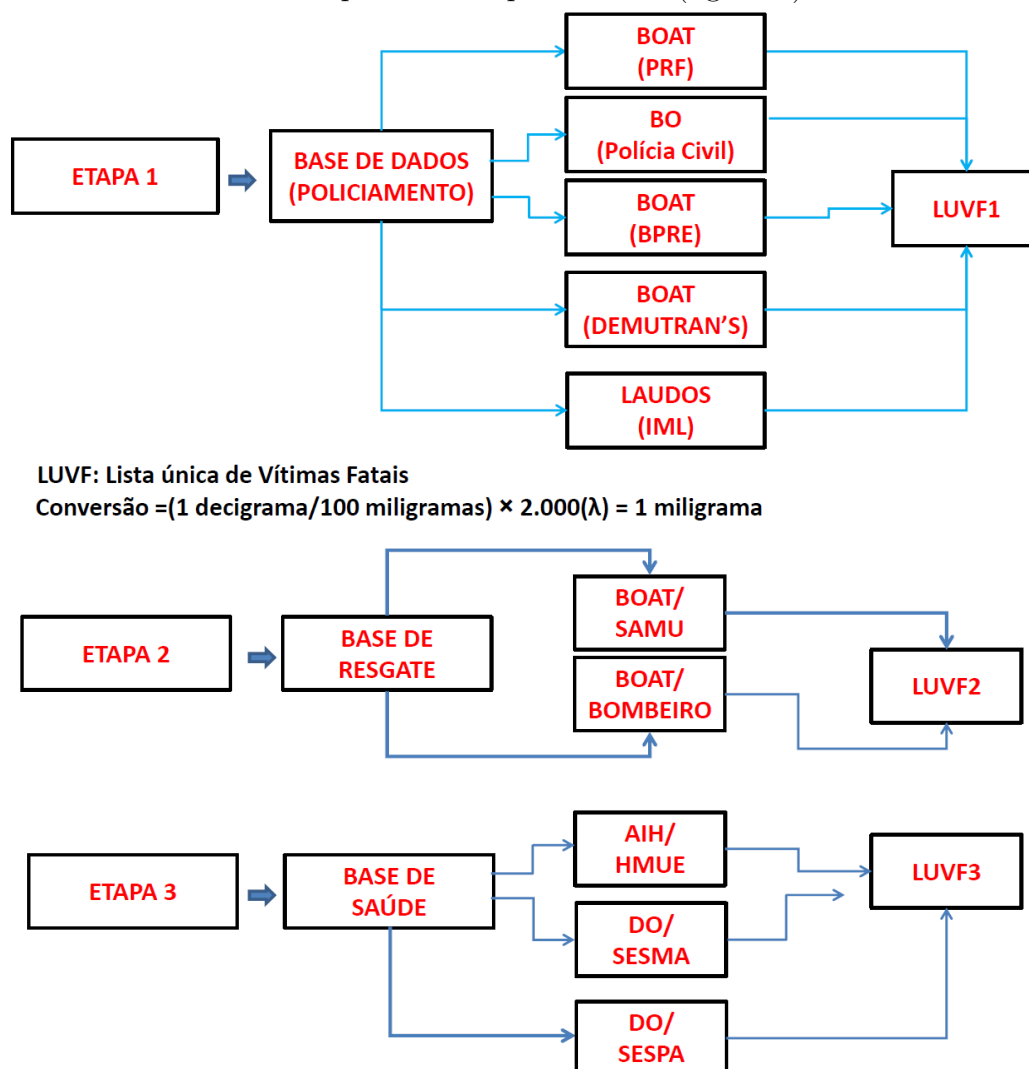
O terceiro passo é utilizar o **Método Probabilístico de Relacionamento de Registros** chamado de Pareamento (*Linkage*), desenvolvido por Fellegi e Sunter (1969).

A principal dificuldade do pareamento é a não existência de um identificador único que permita vincular um Boletim de Ocorrência (BO) de um indivíduo à uma Autorização de Internação Hospitalar (AIH) ou Declaração de Óbito (DO).

Para contornar esse problema, foram definidas as variáveis: **Nome**, **Sexo**, **Data de Nascimento da Vítima** e **Nome da Mãe**, que irão compor a chave de identificação, estimaram-se as probabilidades, por meio de escores, de determinados registros dos bancos de dados dois a dois, pertencentes as mesmas vítimas.

De acordo com a figura 2, as diversas bases de dados foram unidas originando-se um banco de dados integrado de vítimas de sinistros de trânsito a partir da utilização de registros que possuíam as informações definidas na etapa 1, durante o diagnóstico situacional.

Assim, o processo de padronização das variáveis, utilizando o método probabilístico de relacionamento, foi realizado para homogeneizar as variáveis de diferentes bases de dados, visando minimizar erros no processo de pareamento (figura 2).



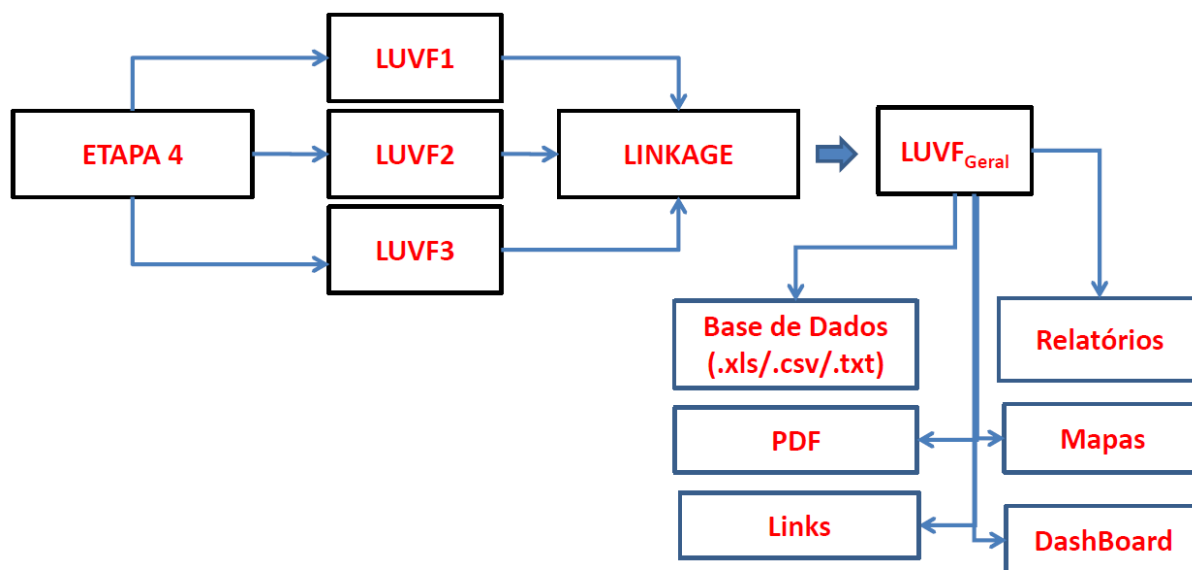
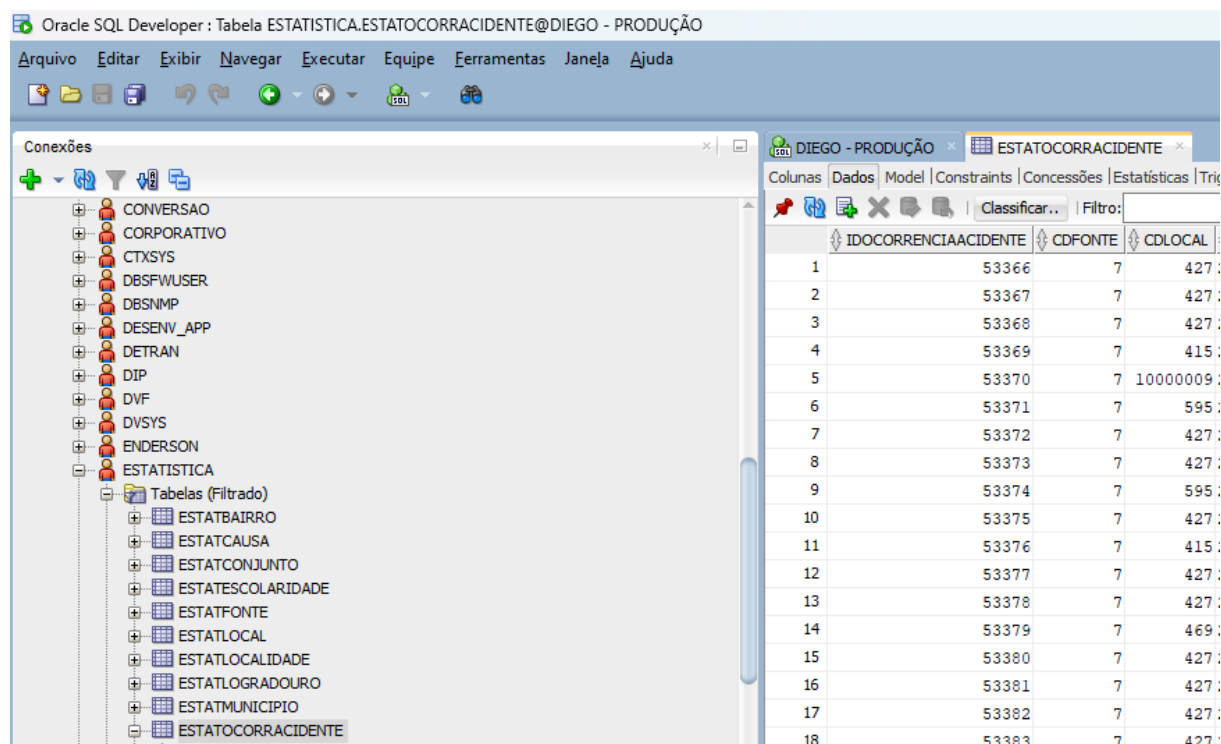


Figura 2: Processo de Gestão de Integração de Múltiplas Base de Dados sobre Sinistros de Trânsito no Estado do Pará, versão março de 2025.

Fonte: DETRAN-PA, 2024

Após o procedimento de pareamento é criado a **Lista Única de Vítimas (LUF)**, na qual as variáveis apresentadas são consideradas mínimas para a sua elaboração. Os dados produzidos são incluídos na base de dados geral do DETRAN-PA.

A figura 3 a seguir apresenta a aba principal das tabelas criadas no ORACLE, para a composição do Banco de Dados após o processo de pareamento, sobre Sinistros de trânsito, hospedado do servidor da Autarquia de Trânsito do Pará.



	IDOCORRENCIAACIDENTE	CDFONTE	CDLOCAL
1	53366	7	427
2	53367	7	427
3	53368	7	427
4	53369	7	415
5	53370	7	10000009
6	53371	7	595
7	53372	7	427
8	53373	7	427
9	53374	7	595
10	53375	7	427
11	53376	7	415
12	53377	7	427
13	53378	7	427
14	53379	7	469
15	53380	7	427
16	53381	7	427
17	53382	7	427
18	53383	7	427

Figura 3: Aba Principal da Base de Dados Geral de Sinistros de Trânsito criado no Oracle Sql Developer, hospedado no data center do DETRAN-PA, versão março de 2025.

Fonte: DETRAN-PA, 2024

O processo de extração dos dados utiliza o pacote *httr* (WICKHAM, 2023) para obter os dados da API da Polícia Civil em formato json, estes dados são manipulados via pacote *jasonlite* (OOMS *et al.*, 2025) e alimentados em um banco ORACLE por meio do pacote *DBI* (WICKHAM e MULLER, 2024).

O primeiro painel lançado para o público externo será o de Sinistros de Trânsito. O menu lateral esquerdo possui algumas informações, como glossário, que explica os conceitos sobre as ocorrências de acidentalidade no trânsito e que apresenta informações sobre a situação socioeconômica, tais como: gênero, idade, escolaridade, estado civil; a dinâmica do evento, com base na condição das vítimas, tipos de sinistros e veículos envolvidos. No que tange à temporalidade, o dia, mês e turno de ocorrência para cada sinistralidade.

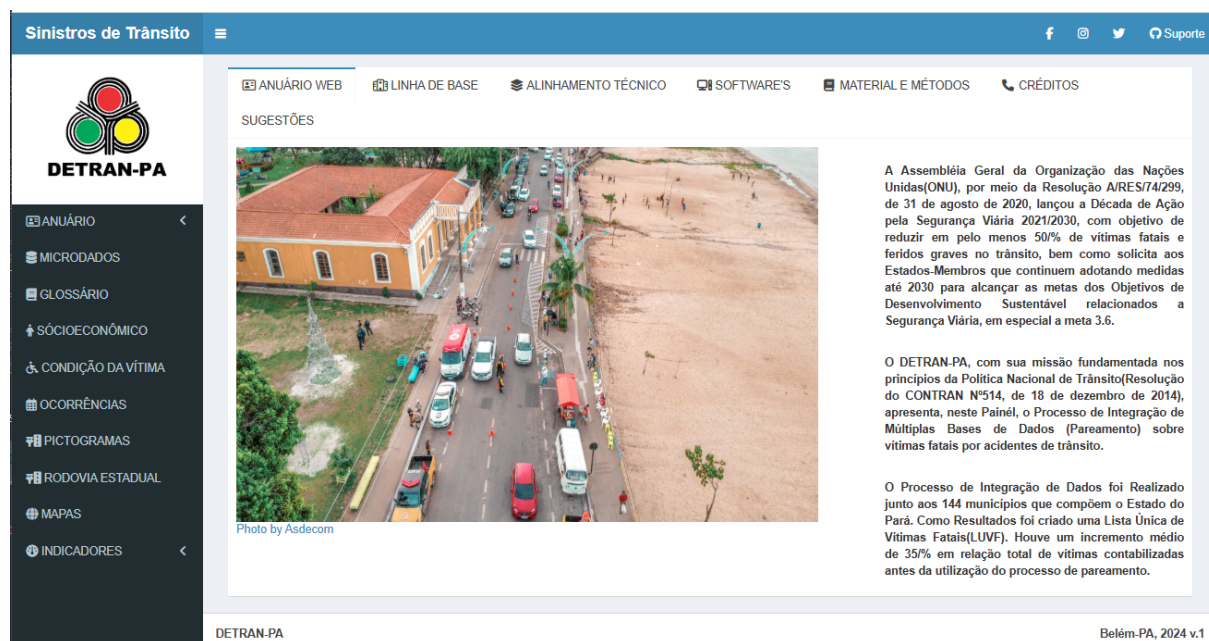


Figura 4: Aba Principal do Pannel de Sinistros de Trânsito criado no framework shiny R, versão março de 2025.

Fonte: DETRAN-PA, 2024

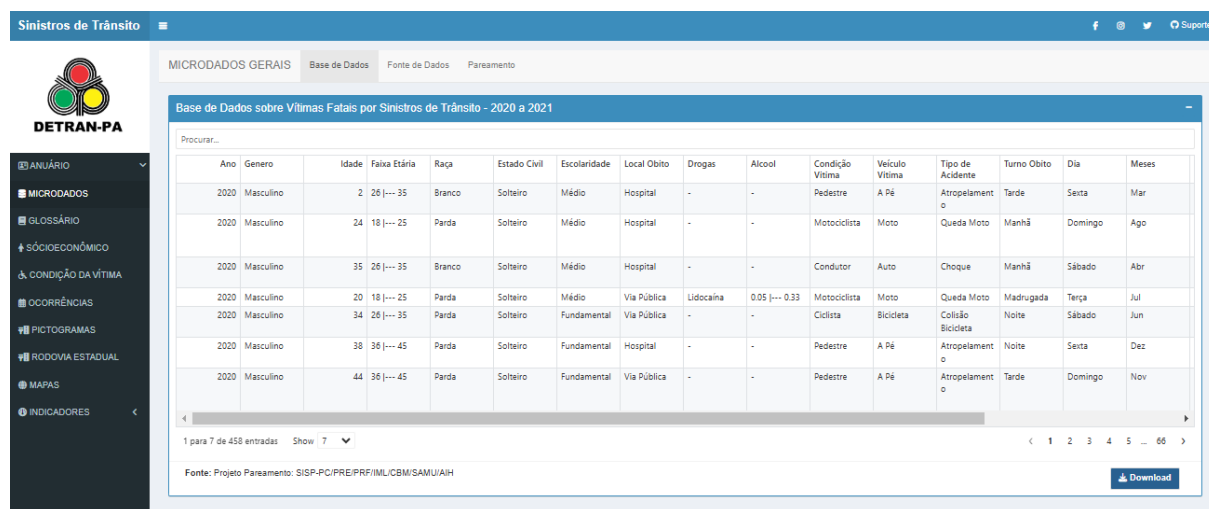


Figura 5: Aba referente aos Microdados do Pannel de Sinistros de Trânsito criado no framework shiny R, versão março de 2025.

Fonte: DETRAN-PA, 2024

A aba de distribuição espacial (figura 6) foi construída de modo a permitir que os usuários consigam visualizar o comportamento no espaço dos diferentes sinistros de trânsito, incluídos no painel por meio de mapas coropléticos. Os filtros permitem delimitar o período (ano) e o nível de divisão espacial entre municípios, e os tipos de região de integração.

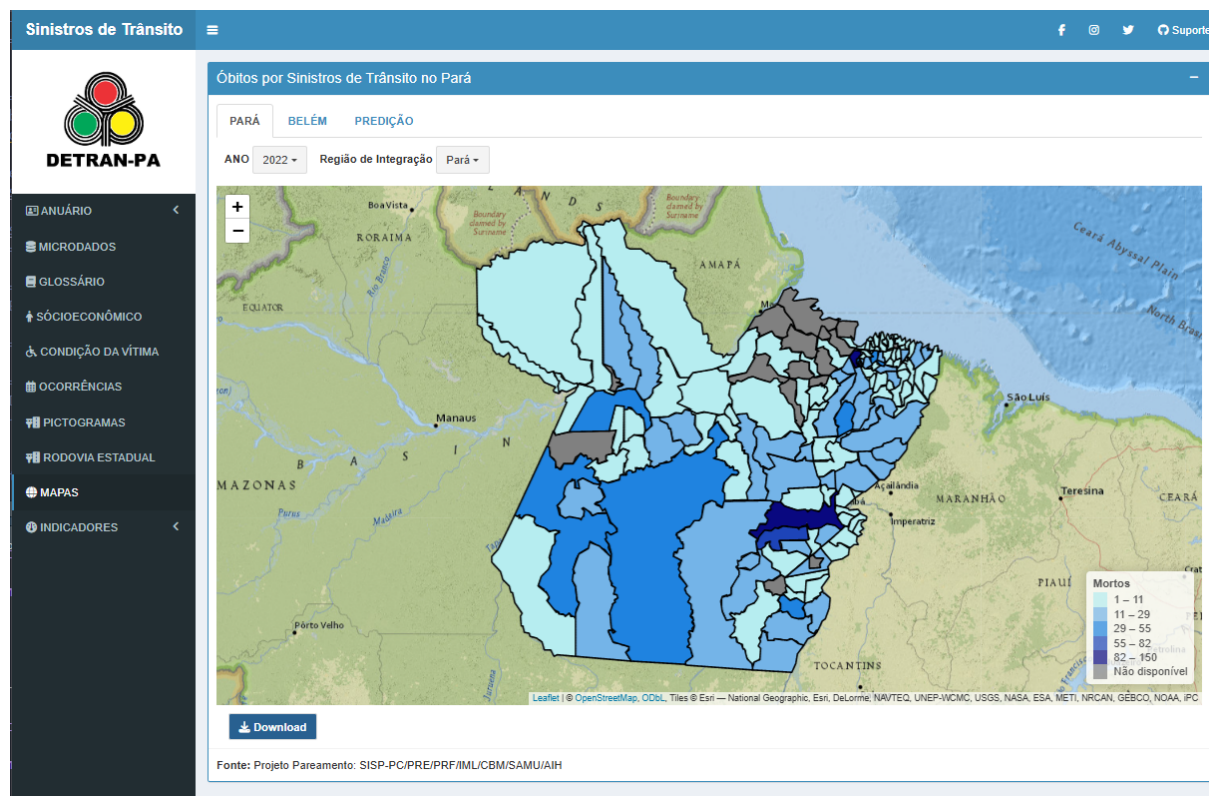


Figura 6: Aba referente a distribuição Espacial do Painel de Sinistros de Trânsito criado no framework shiny R, versão março de 2025.

Fonte: DETRAN-PA, 2024

O último item do menu é referente aos indicadores, detalhando sobre as principais taxas de mortalidade por sinistros de trânsito. A figura 7 ilustra as taxas de mortalidade por 100 mil habitantes na condição geral e motociclista.

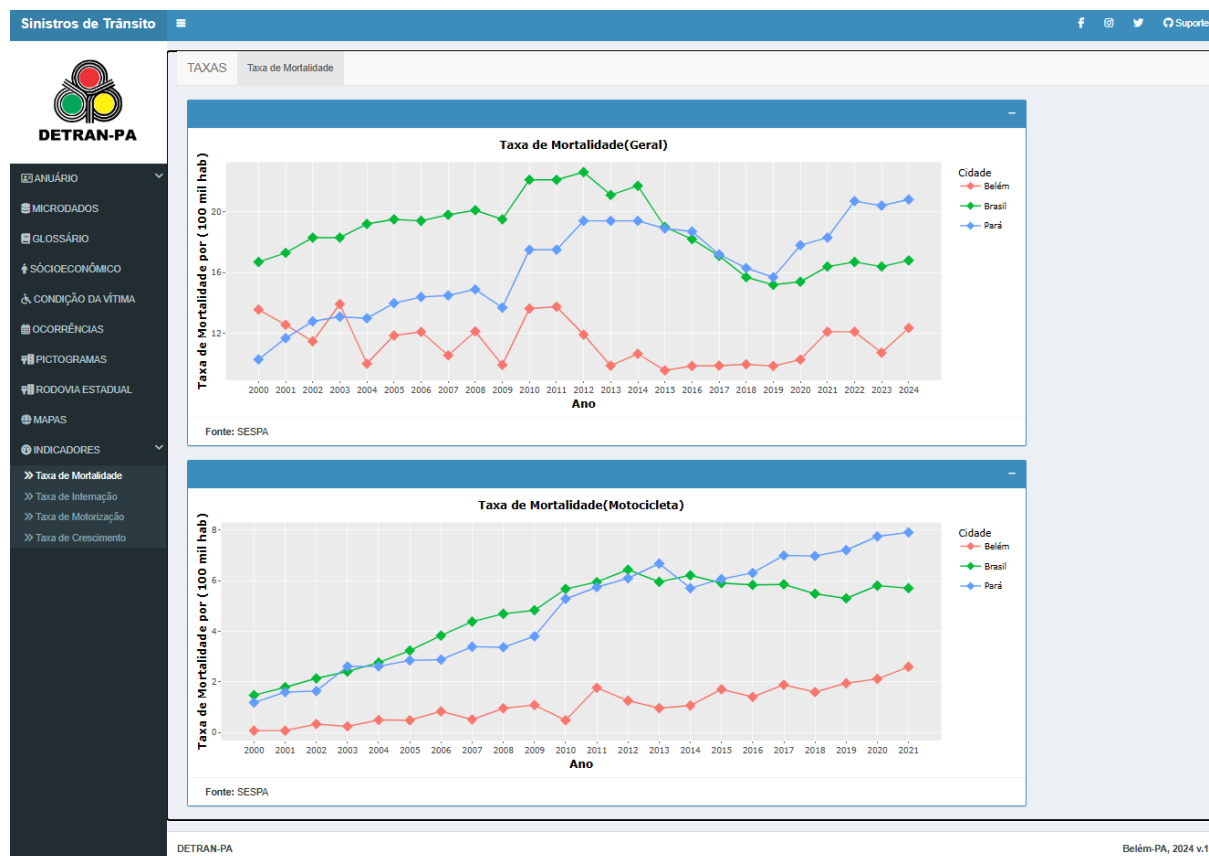


Figura 7: Aba referente aos Indicadores (taxas) do Painel de Sinistros de Trânsito criado no framework shiny R, versão março de 2025.

Fonte: DETRAN-PA, 2024

Conclusão

Ao longo dos anos, em sua missão de dar mais transparência aos dados de segurança viária do Pará, o Departamento de Trânsito do Estado do Pará vem desenvolvendo diversos produtos para tornar estas informações mais acessíveis para a população. O dashboard sobre sinistros de trânsito surge como uma alternativa de modernização, porém baseado em *software* livre.

Para trabalhos futuros, sugere-se um modelo de previsão de sinistros de trânsito usando um modelo Autor regressivo Integrado de Médias Móveis Sazonal (SARIMA) com o pacote *fable* tendo como critérios de escolha os valores de Erro Percentual Absoluto Médio (MAPE), para geração de estimativas de Previsão para cinco anos à frente.

Referências

COENE, John. **echarts4r: Create Interactive Graphs with 'Echarts JavaScript' Version 5**. R package version 0.4.5. 2023. Disponível em: <https://cran.r-project.org/web/packages/echarts4r>.

CHANG, Winston. CHENG, Joe. ALLAIRE, JJ. SIEVER T, Carson. SCHLOERKE, Barret. XIE, Yihui, ALLEN, Jeff. MCPHERSON, Jonathan. DIPER T, Alan. BORGES, Barbara. **Shiny: web application framework for R**. R package version 1.10: 2024. Disponível em:

CHENG, Joe. KARAMBELKAR, Bhaskar. XIE, Yuhui. **leaflet: create interactive web maps with the javascript leaflet library**. R package version 2.2.2: 2024. Disponível em: <https://cran.r-project.org/web/packages/leaflet>.

DETRAN-PA. *Relatório de Gestão 2024*. Belém: Departamento Estadual de Trânsito do Pará, 2025. Disponível em: [<https://www.detran.pa.gov.br/files/transparencia/nucleoPlanejamento/RelatoriodeGestao2024.pdf>]. Acesso em: 22 abr. 2025.

FELLEGI, I. P.; SUNTER, A. B. A theory for record linkage. Journal of the American Statistical Association, v. 64, n. 328, p. 1183-1210, 1969.

GRANJON, David. **bs4Dash: a bootstrap4 version of shinydashboard**. R package version 2.1.0: 2022. Disponível em: <https://cran.r-project.org/web/packages/bs4Dash>.

OOMS, Jeroen. **jsonlite: a Simple and Robust JSON Parser and Generator dor R**. R package version 1.9.1: 2025. Disponível em: <https://arxiv.org/abs/1403.2805>.

POSIT TEAM. RStudio: Integrated Development Environment for R. Posit. Software, PBC, Boston, MA. 2024. URL <http://www.posit.com>.

R CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2024.

R Special Interest Group on Databases (R-SIG-DB), Wickham H, Müller K. **DBI: R Database Interface**. R package version 1.2.3, 2024. Disponível em: <https://cran.r-project.org/web/packages/DBI>.

XIE, Yihui. CHENG, Joe. TAN, Xianying. **DT: a wrapper of the javascript library datatables**. R package version 0.1.9: 2021. Disponível em: <https://cran.r-project.org/web/packages/DT>.

WICKHAM, H. **httr: Tools for Working with urls and http**. R package version 1.4.1. 2023. Disponível em: <https://cran.r-project.org/web/packages/httr>.

Anexo

Foi desenvolvido um script na linguagem de Programação R, para gerar o dashboard, armazenado no repositório GitHub que pode ser acessado no seguinte link: https://github.com/MarioDhiego/Dashboard_Modulo_Sinistro_2024