

A DASHBOARD FOR DECISION-MAKING CONSIDERING OPEN DATA FROM FEDERAL HIGHWAYS

Rodrigo Gabriel de Miranda

rgabrieldemiranda@gmail.com
Santa Catarina Federal University -
UFSC, Florianópolis, SC, Brazil.

Andréa Cristina Konrath

andreack@gmail.com
Santa Catarina Federal University -
UFSC, Florianópolis, SC, Brazil.

Maicon Rafael Zатели

maicon.zатели@ufsc.br
Santa Catarina Federal University -
UFSC, Florianópolis, SC, Brazil.

Yohana Taise Hoffmann

yohana.thc@gmail.com
Santa Catarina Federal University -
UFSC, Florianópolis, SC, Brazil.

Silvio Aparecido da Silva

silvioest@gmail.com
Santa Catarina Federal University -
UFSC, Florianópolis, SC, Brazil.

ABSTRACT

Highlights: The law of access to information guarantees that data is made available to the population. This data is in raw format and has no real use. Traffic accident records can be more informative. This requires the development of web technologies based on data science design concepts. **Objective:** Develop an informative dashboard with a user-friendly interface and web technology to make useful and simplified data available to the population. **Methodology:** Data regarding accidents on Brazilian federal roads, registered by the Federal Highway Patrol and made available as open data, were collected. Based on the principles established by the Open Government Data Initiative and through the concepts of data science projects associated with the R programming language, a dashboard was developed for interactive consultation of the accident data on federal roads. **Results:** The dashboard developed can provide information with a friendly interface, relating the recorded accidents with environmental parameters and road conditions, such as weather conditions, characteristics of the road sections, region, temporality, and the driver's sex and age. The information is presented in the dashboard through graphs, tables, and interactive maps, with the possibility to search accidents by combining filters and allowing access to direct information from accident statistics. **Practical Implications:** The dashboard developed, by providing clear and simplified useful data, serves as a support for strategies to prevent accidents on highways. The Federal Highway Patrol Department can use these data to reduce the inconvenience caused by traffic accidents. **Originality:** This dashboard model, focused specifically on traffic accident information management, brings innovation to processing and analyzing data from federal highways. Therefore, it aims to contribute to the public management of accidents in a simple manner by providing simplified information to the population.

Keywords: Accessibility; Digital Inclusion; Data Analysis; Traffic Accidents.

INTRODUCTION

Information and Communication Technologies (ICTs) are increasingly present and have contributed to changes in various aspects of the population, especially those related to access to information available in databases. Society has been demanding that governmental organizations offer better quality services, especially regarding transparency in the management of public resources. As far as public administration is concerned, the sector works with a large volume of data daily; however, this data is not always properly available to the population.

This issue has generated movements in administrative thought, such as Open Government, which, according to Veljković *et al.* (2014), began with measures from the U.S. government to increase the transparency mechanisms. The consolidation of Open Government came with the official declaration of the Open Government Partnership (OGP), of which Brazil and seven other countries have been members since 2011 (Open Government Partnership, 2011).

Transparency mechanisms are related to the availability of open data, through which it is possible to know the government's actions in a more detailed way (Silva *et al.*, 2014). According to the definition of the Open Knowledge Foundation (OKF), "data are open when anyone can freely use, reuse, and redistribute them, and they are subject, at most, to requiring crediting their authorship and sharing under the same license" (OKF, 2020). According to Attard *et al.* (2015), different global government open data actions have emerged in the last decade aimed at data transparency and reuse.

Brazilian Law No. 12,527 was created on November 18, 2011 (Brazil, 2011) and is known as the "Access to Information Law." This law allowed the availability of government data, but in many situations only in its raw form, such as records in tables or an Application Programming Interface (API), requiring some sort of processing so it could become information.

Among the public data opened by the Brazilian government are the bases related to traffic accidents, which are available on the website of the Federal Highway Patrol Department (DPRF, 2019a). However, such data are currently only available in raw form as tables. Thus, when accessing this database, the information is not laid out straightforwardly to contribute to answering questions such as: (i) in which month of the year more accidents occur on a given highway; (ii) on which stretches of a given highway there are more accidents; (iii) how is the distribution of accidents on a given highway, considering the lane type (single, double, multiple); (iv) how is the distribution of accidents on a given highway, considering the severity (without victims, with

injured victims, with fatal victims); (v) What are the principal causes of accidents on a given highway? (vi) What is the most common age range of people involved in accidents in a given state? Moreover, most of the time, exploring this type of data becomes harder as it requires technical knowledge in the area and demands more advanced analysis techniques (Braunschweig *et al.*, 2012).

Road traffic crashes are noted as the leading cause of death among young people worldwide, ranking first in the 15–29 age group and second in the 5–14 age group. Moreover, it is estimated that in 2016, about 1.35 million people died in traffic accidents (WHO, 2018). Considering the proportion of these numbers, this theme has been addressed as a significant public health issue. In Brazil, a series of actions implemented by the federal, state, and municipal governments (educational campaigns, which include the Yellow May and the National Traffic Week; the intensification of enforcement, such as the Dry Law blitz; the improvement of vehicle safety; and traffic engineering measures, such as the modernization of highways) have reduced the annual number of deaths from traffic accidents. Between 2015 and 2019, there was a 7% annual drop in the number of deaths, which, according to the Computer Department of the Unified Health System (DataSUS), is a reduction from 43,000 to 30,000 deaths yearly (Government of Brazil, 2020).

Traffic accidents cost about 3% of most countries' gross domestic product (GDP) (Pan American Health Organization, 2019). Considering both direct and indirect costs for the year 2014, traffic accidents generated a loss of around R\$ 40 billion to Brazil, and of these, R\$ 12.3 billion stems from accidents on federal highways (Institute of Applied Economic Research, 2015). Thus, it is possible to realize that the high incidence of traffic accidents directly affects the national economy (Tischer, 2019).

Based on this issue, this article aims to present a tool (in dashboard format) that transforms the open data concerning traffic accidents provided by the DPRF into information available to the general population and also to the public administration, especially the DPRF itself. On the one hand, by accessing the dashboard, the general population can obtain more direct answers to questions such as those listed above. On the other hand, the dashboard presents information that has already been worked out and is more structured than the database currently available on the DPRF website. It also serves as a basis for prevention strategies the DPRF employs to reduce the inconvenience caused by traffic accidents, such as deciding on a given stretch of road to be more heavily inspected by police officers during a given period of the year, day of the week, or time of day, or on the location of a new service station or new electronic surveillance. The highway administration (public or private) would also have more

information to, for example, decide about infrastructure works, signaling stretches, and support for highway users.

The dashboard developed for the DPRF website makes some of the following functionalities available: (i) various graphs, which consider the temporal, geographic, conditioning, severity, and involved characteristics; (ii) interactive maps that allow visualizing accidents by state and highway; (iii) search for accidents by combining some filters by the municipality, state, highway, date of occurrence, etc. These will be presented in the results and discussions. At the end of this article, the importance of the Access to Information Law of 2011 and the availability of open data are highlighted, along with the conclusions obtained and directions for future work.

Open Data

With the legislation on transparency in the public sector, the open data movement has increased in Brazil. According to Harrison *et al.* (2012), open data contributes to greater social control, the strengthening of democracy and citizenship, and an improvement in public administration. The development of open data actions, along with the innovation and diffusion of new digital devices, allow greater accessibility to data and the ease of developing tools for processing, storage, manipulation, analysis, and distribution of this vast data volume (Boulton, 2014; Boulton *et al.*, 2011; Braunschweig *et al.*, 2012). In addition, the Federal Audit Court (TCU, 2015) presents five main reasons for opening data: (i) transparency in public management; (ii) contribution of society with innovative services to citizens; (iii) improvement in the quality of government data; (iv) enabling new business; and (v) obligation by law.

Pinho and Silva (2019) argue that government information and open data differ in prerequisites, as the open format standard, elaborated by a series of requirements, emphasizes raw data formatted so that it can be easily worked on and analyzed by third parties. Thus, on the one hand, open data is information not previously interpreted, but open data users can directly access this raw data and customize it according to their needs (Braunschweig *et al.*, 2012). On the other hand, government information is usually published in formats that make it difficult to access and use, preventing it from being combined and aggregated with other data (Diniz, 2010). Thus, studies to elaborate mechanisms for data integration are necessary, aiming at their interconnection and reuse, as well as studies to explore their economic potential (Albano *et al.*, 2019).

According to the Brazilian Open Data Portal (PBDA, 2020), eight principles govern open data, namely:

- **Complete:** All public data are made available. Data are electronically recorded information, including, but not limited to, documents, databases, transcripts, and audiovisual recordings. Public data are data that are not subject to valid privacy, security, or access control limitations governed by statute;
- **Primary:** Data are published in the form collected at the source, with the finest possible granularity, and not in aggregated or transformed form;
- **Current:** Data are made available as quickly as they are needed to preserve their value;
- **Accessible:** Data are made available to the widest possible audience and for the widest possible range of purposes;
- **Machine processable:** The data are reasonably structured to enable automated processing;
- **Non-discriminatory access:** The data is available to all without requiring identification or registration;
- **Non-proprietary formats:** The data is available in a format over which no single entity has exclusive control;
- **License-free:** Data are not subject to copyright, trademark, patent, or trade secret regulations. Reasonable restrictions on privacy, security, and access control may be permitted, provided they are in the form regulated by statute.

In addition to these, Open Government Data (2007) lists seven additional principles that are not included in the original principles for open government data but which could also be considered:

- **Online and free:** The data must be available on the Internet free of charge, and it must be possible to find it;
- **Permanent:** *The data must be available indefinitely, in a stable location on the Internet, and a stable data format for as long as possible;*
- **Trustworthy:** Published content should be digitally signed or include an attestation of the publication date, authenticity, and integrity. Digital signatures help the public validate the origin of the data found and ensure that it has not been modified since publication;

- **Openness presumption:** The openness presumption is based on laws such as the Freedom of Information Act, procedures that include records management, and tools such as data catalogs. Defining the standard as open means that the government and parties acting on its behalf will proactively make public information available and placed within the public’s reach (online);
- **Documented:** Government websites should make sufficient information available to users by documenting the data’s format and meaning, making it useful;
- **Safe to open:** Government agencies that publish data online should always seek to publish using data formats that do not include executable content, as such content within documents poses a security risk to data users because it may be malware;
- **Designed with public contribution:** The public is in the best position to define which information technologies will best suit the applications they want to create for themselves. Public input is thus crucial to disseminating information in a valuable way.

These principles were used to analyze the DPRF open data portal (2019a, 2019b), and their analysis results are detailed in the next section.

MATERIAL AND METHODS

Data collection and analysis

According to Gil (2017), this research is considered applied in nature and quantitative in form because its results are used to solve real problems, seeking to confirm through numbers the knowledge generated in its application. The dataset is related to traffic accidents on federal highways in Brazil from 2007 to 2019, obtained from the DPRF website (2019a). This dataset includes several variables, such as cause of accident, weekday of occurrence, municipality of occurrence, and classification. More details about the variables in the database can be found in DPRF (2019b).

Considering the eight principles cited in the previous subsection (PBDA, 2020), the DPRF website can be classified for each of them, as shown in **Chart 1**.

Chart 1. Evaluation of access to information on the DPRF website

Principle	Evaluation
Complete	Yes, it has data on the characteristics of the accidents and the injured.
Primary	Yes, each line of the table is a record of an accident in Brazil.
Current	Partially. The data is updated quarterly.
Accessible	Yes.
Machine Processable	Partially. The data is in table format but has some inconsistencies, requiring processing for cleaning and normalization.
Non-discriminatory access	Yes, no identification is required to access the data.
Non-proprietary format	No. The data is in a compressed .rar format extension, which is proprietary.
License-free	Yes.

Source: Prepared by the authors (2020).

Although the DPRF website partially meets the eight principles of open data, it is not very useful for accident research because the data needs to be processed to generate any information. For example, if the user wants to know which highway had the highest number of accidents in Santa Catarina (SC) in 2018, this information will not be available, requiring knowledge of some programming language for data aggregation.

Chart 2 shows the seven additional principles analyzed regarding the DPRF website for each of them.

Chart 2. Additional principles evaluated on the DPRF website

Principle	Evaluation
On-line & gratuito	Yes. The data is available online and free of charge and can be accessed at https://portal.prf.gov.br/index.php/dados-abertos .
Online & free	Yes, the data are available on the DPRF website, are compressed (zipped) to reduce their actual size, and are in .csv (comma-separated values) format.
Permanent	The data are not digitally signed, nor do they have certification of publication date, authenticity, or integrity.
Trustworthy	Yes, the DPRF updates and makes the data available periodically on its website and proactively, complying with the legislation on transparency in the public sector.
Openness presumption	Yes, data dictionaries are provided to facilitate understanding of the data files’ contents.
Documented	Yes, data are made available compressed in .zip, .rar, or .csv format, and there is no executable content.
Secure to open	No evidence was identified that there is a public contribution to the open data formats and content made available by the DPRF.

Source: Prepared by the authors (2020).

To improve access to this type of information, a data science design approach was used, utilizing the R programming language, as described in the next section.

Data Science Project

Figure 1 shows a typical data science project. This project was developed using the R programming language. The R software, which is an open source project, is a language and environment for statistical computing (R Core Team, 2020). Its programming is done by typing command lines, meaning that users using this program must have some programming language knowledge. Since not everyone has this knowledge, the Shiny package (Chang and Ribeiro, 2018), a framework for developing web applications, is an alternative. By using Shiny, it is possible to replace function programming in R in interactive web applications. In addition to Shiny, some more packages (R currently has over 16,045) useful for the project are described in the figure.

Description of the steps of a data science project, according to Wickham and Grolemund (2017):

- Define the problem: Defining the problem helps identify the data we need to answer the questions raised;
- Import data: Data import can happen by accessing a database, downloading from some website, or scraping data from the internet. Packages recommended for the R software: readr (Cooper, 2017), readxl (Wickham and Bryan, 2019), httr (Wickham, 2020), and dbi (Wickham and Müller, 2019);

- Fixing the data: Consists of removing inconsistencies and normalizing them. Package recommended for the R software: tidyr (Wickham and Henry, 2020);
- Transform: It consists of modifying the data into tables, where each row is an observation and each column a variable. The transformation of raw data into information takes place in this step. Packages recommended for the R software: dplyr (Wickham *et al.*, 2020b) and purrr (Henry and Wickham, 2020);
- Visualize: The charts generated in the previous step can be transformed into graphs. The transformation of raw data into information takes place in this step. Recommended packages for the R software: ggplot2 (Wickham *et al.*, 2020a) and plotly (Sievert *et al.*, 2020);
- Modeling: Statistical models can be generated to aid in the production of information. Packages recommended for the R software: broom (Robinson *et al.*, 2020) and sweep (Dancho and Vaughan, 2019);
- Communicate: Once it is defined which information is relevant, i.e., which graphs and tables to show, the reports or dashboards are produced. Packages recommended for the R software: rmarkdown (Allaire *et al.*, 2020) and shiny (Chang *et al.*, 2020).

Thus, after the steps of the data science project and their respective packages for the R software, it is possible to develop the dashboard for the DPRF website to contribute to the treatment of the database and access to public information.

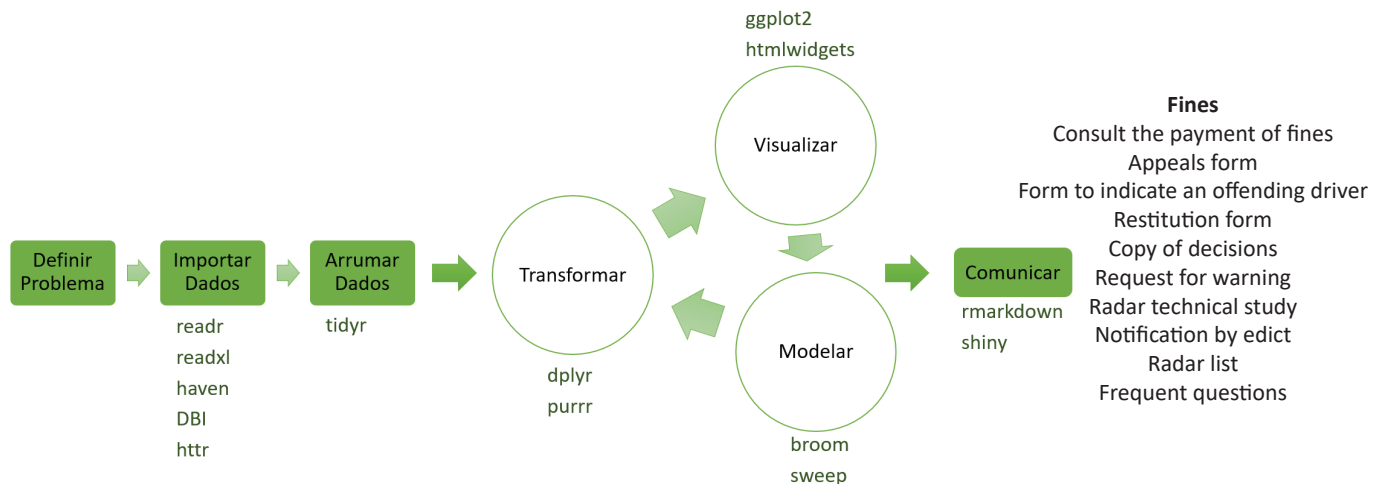


Figure 1. Typical data science project
 Source: Adapted from Wickham and Grolemund (2017).

RESULTS AND DISCUSSION

Following the methodology described, including the phases of the data science project, a dashboard was developed for the DPRF website, and its database can be seen as follows, as shown in **Figure 2**.

On the site in **Figure 4**, there are only links to download the accident data. If, for instance, there is a need to make an analysis over the years, it is necessary to download all the years, open the compressed files in .rar format, and use a program to join the tables, “clean” the data, and transform (make some aggregation). Only then is it possible to create graphs and do analyses. These procedures are quite distant from people because they require specific knowledge in computing and statistics. In this sense, a dashboard was developed to assist in accessing and visualizing the data, which can be accessed at the following link: <https://r-nnq.shinyapps.io/dashboard>.



Figure 2. Access to data from the DPRF website.

Source: Department of Federal Highway Police (2019a).
Federal Highway Police

Ministry of Justice and Public Security

Accidents

Grouped by occurrence

2019 - Jan to Jun; 2018; 2017; 2016; 2015; 2014; 2013; 2012; 2011; 2010; 2009; 2008; 2007

Grouped by person

2019 - Jan to Jun; 2018; 2017; 2016; 2015; 2014; 2013; 2012; 2011; 2010; 2009; 2008; 2007

Grouped by person - All causes and types of accidents (from 2017)

2019; 2018; 2017

In the developed dashboard (**Figures 5 and 7–20**), it is possible to perform different queries. The left-side menu has information, such as Search by accident and Description of the Roads. When you select “Search,” search filter ele-

ments will open, such as: choose a state; choose a BR highway; choose a km range; filter by date. After selecting each element, go to “Apply Search” (**Figure 3**).

The dashboard will show the statistical data of the query performed: it presents general data on the number of accidents, fatalities, injured victims, and the state/BR road. The search features are the road map, temporal, geographic, conditioning, severity, and people involved.

To show the potential of the developed dashboard, which can contribute to decision-making for implementing public policies, the BR 101 and BR 470, which cross the state of Santa Catarina, were selected. According to the National Department of Transportation Infrastructure (DNIT, 2020), BR 101 is considered a longitudinal highway that cuts the country in a north-south direction, whereas BR 470 is considered a connecting highway, linking federal highways or at least one federal highway to cities or relevant points. BR 470, in Santa Catarina, involves several municipalities in its stretch, which concentrate industrial centers in the most diverse areas, as well as a port and airport, thus enabling the flow of their products.

Regarding the number of accidents on these two highways, the study by Peña *et al.* (2008) shows the evolution in the number of accidents, as seen in **Figure 4**.

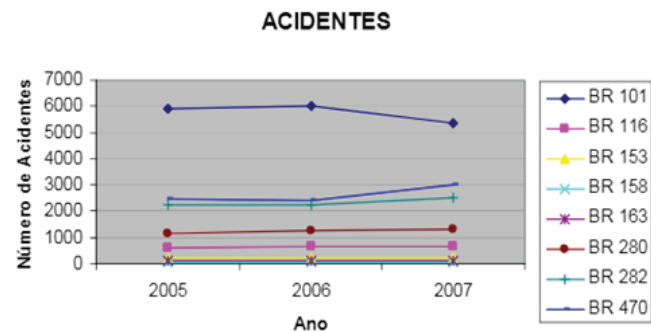


Figure 4. Graph with the number of accidents for the years 2005, 2006, and 2007.

Source: Peña *et al.* (2008, p. 14).

According to the aforementioned authors, BR 101 has a bigger difference regarding the number of accidents compared to the other highways in Santa Catarina State. According to Peña *et al.* (2008), BR 470 stands out for the increasing number of accidents per year. In the study conducted by Possamai (2017), in which the DPRF database was used, 13,270 accidents were identified on BR 470 in Santa Catarina from 2012 to 2017. The author characterized this highway as having the highest morbidity and mortality rates per road kilometer. Thus, the choice of these two highways is justified.

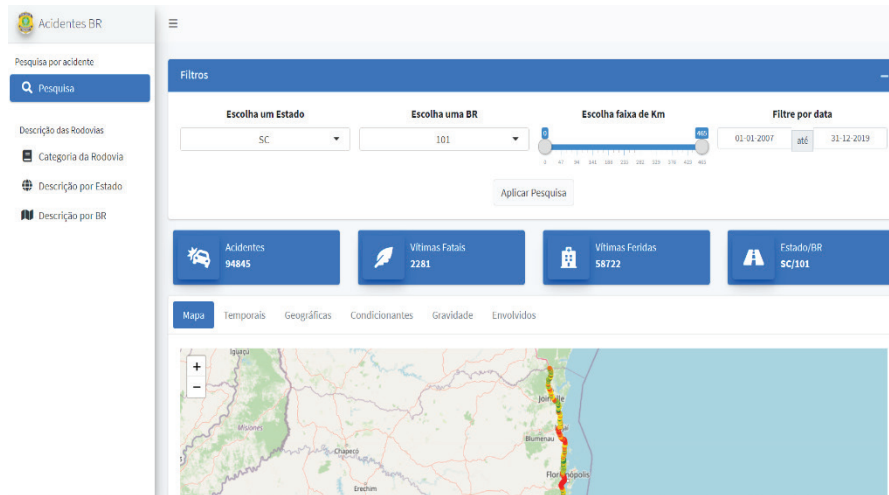


Figure 3. Dashboard search tool

Source: Prepared by the authors (2020).

Next, we present some queries in the dashboard that can contribute to the decision-making process for implementing public policies.

(i) In which month of the year do more accidents occur on a given highway?

This is a question to which the general public can find the answer by viewing the graphs when entering the dashboard. The query to find out in which month of the year there are more accidents on a given highway requires some steps and filters. In this case, you choose the state of Santa Catarina, BR 101, and the period from January 1, 2007 to December 31, 2019. Then, press “apply search,” and the “Map of Accidents” and the characteristics: temporal, geographical, conditioning, severity, and involved will appear. In this query, we selected the characteristic “temporal” and the variable “month of the year” (**Figure 5**). The same procedure is performed for the consultation of BR 470 (**Figure 6**).

Thus, we can quickly discover which month presents the highest number of accidents. We can see that December (BR 101: 9,247; BR 470: 3,252) and January (BR 101: 8,919) have the highest number of accidents. This increase in accidents during this period may be related to the end-of-year recess. BR 101 cuts through the SC coast, and BR 470 is one of the roads that connect the inland cities to the coast. Therefore, the flow of vehicles from various parts of the country to the state and within the state itself, is greater than the traffic during the other months of the year. One of the characteristics of the state of Santa Catarina is tourism, which receives many visitors during the summer months. According to the Santa Catarina government news portal, the 2019/2020 season had an expectation of 5 million tourists. This information is relevant for thinking about

different strategies for public safety and, consequently, the reduction of traffic accidents (for more information, see Gorges, 2019).

(ii) On which stretches of a given highway do accidents occur the most?

The map reveals some information when verifying the stretches where more accidents occur. However, we can perform the query by municipality and km. The procedure is the same as in the previous topic; however, we select the “geographical” characteristic, and “municipality” is the chosen variable. And then, in the new query, the chosen variable is “Km.”

Figure 7 illustrates the number of accidents on each stretch of BR 101 from 2007 to 2019, and visually, we can see that the highest concentration of accidents is in the stretch corresponding to Greater Florianópolis, i.e., from kilometer 200 to 220. This data corroborates **Figure 8**, when we check the accidents per municipality. São José has 17,428 accidents, followed by Palhoça with 11,499 accidents.

Figure 9 illustrates the number of accidents on each stretch of BR 470 from 2007 to 2019, and visually we can see the stretch with the highest concentration of accidents, corresponding to the Médio Vale do Itajaí region, from kilometer 40 to 60. We can confirm this data by checking the accidents per municipality, according to **Figure 10**. Blumenau ranks first with 7,610 accidents, followed by Indaial with 4,664 accidents.

In this way, the places with the highest frequency of traffic accidents may be related to several causes, such as a higher flow of vehicles because they are metropolitan regions (central cities that influence the others economically,

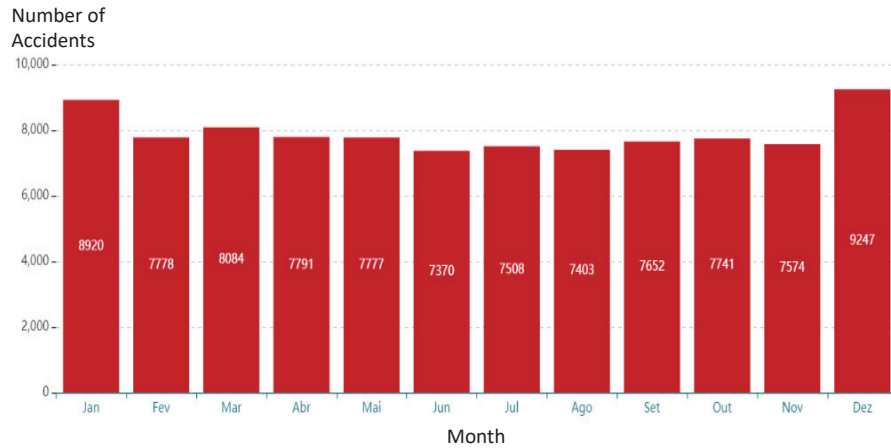


Figure 5. Month with more accidents on BR 101 in Santa Catarina
Source: Prepared by the authors (2020).

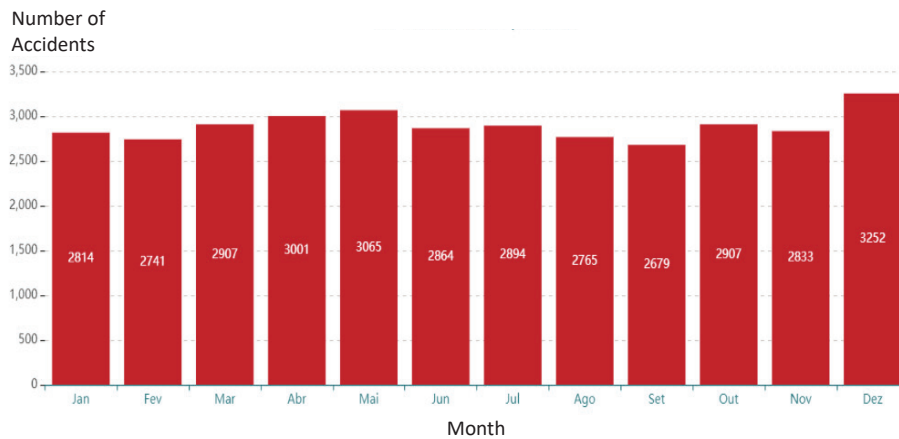


Figure 6. Month with more accidents on BR 101 in Santa Catarina
Source: Prepared by the authors (2020).

socially, and politically), a lack of signaling, and the number of accesses, among others.

(iii) What is the distribution of accidents on a given highway, considering the lane type (single, double, or multiple)?

To identify the type of road with the highest incidence of accidents, the “geographic” characteristic and the “type of road” variable must be selected in the query. This step-by-step must be performed on both BR 101 and BR 470. We can see (Figure 11) that most accidents on the BR 101 occur on a dual carriageway (66.51%). On the other hand, on BR 470, the vast majority of accidents occur on a single lane, corresponding to 98.7%.

It is worth pointing out that the reality of the two BRs is different: BR 101, in the state of SC, has undergone a duplication process, and currently its route in the state is comple-

tely duplicated. BR 470, on the other hand, is single-lane for the most part, and currently there is a stretch that is in the process of duplication (between the towns of Navegantes and Indaial, approximately 70 kilometers; for more information, see Vieira, 2020), but there are tensions and economic and political obstacles that prevent the works from advancing.

(iv) How is the distribution of accidents on a given highway, considering the severity (without victims, with injured victims, and with fatal victims)?

Regarding the severity of traffic accidents, it is possible to verify this by selecting the characteristic “severity” (Figure 12). We performed the same procedure for BR 101 and BR 470, and even though the two roads have particular characteristics, we found similarities. A little more than half of the accidents have no victims (BR 101, with 53.16%, and BR 470, with 50.34%), followed by accidents with injured vic-

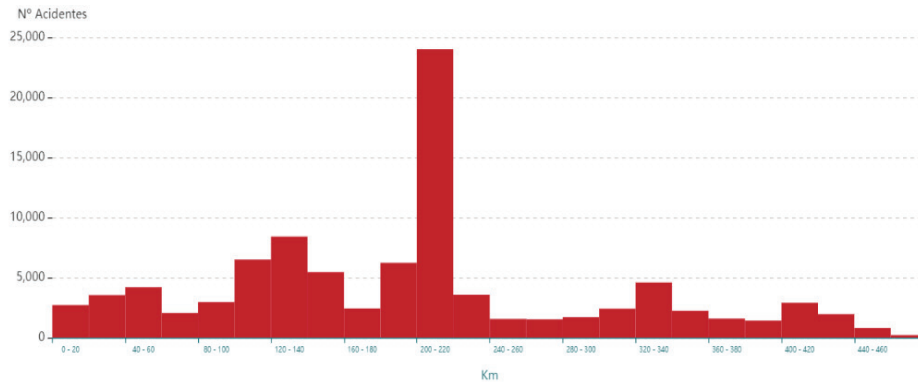


Figure 7. The stretch with the most accidents on BR 101, in Santa Catarina
 Source: Prepared by the authors (2020).

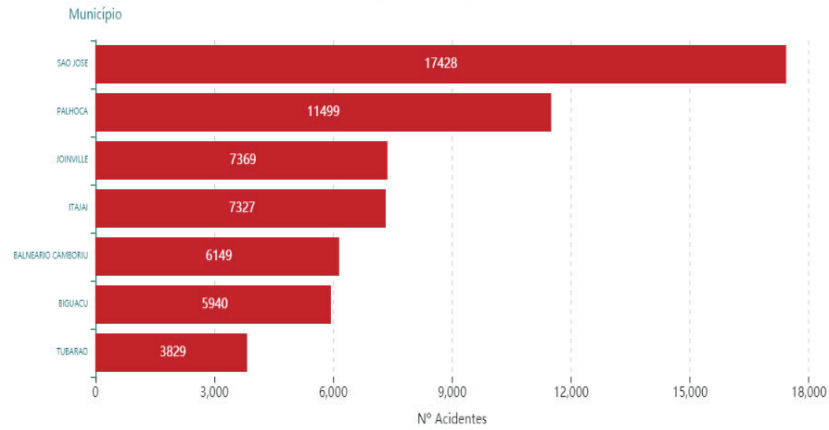


Figure 8. Accidents per municipality on BR 101 in Santa Catarina
 Source: Prepared by the authors (2020).

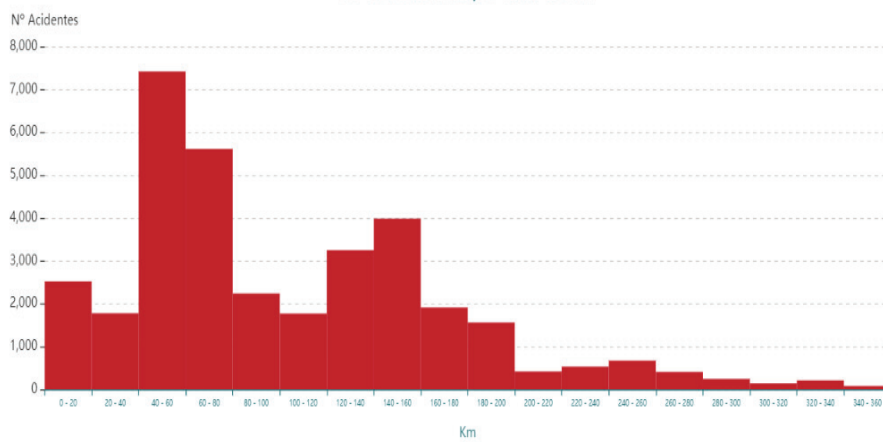


Figure 9. The stretch with more accidents on BR 470 in Santa Catarina
 Source: Prepared by the authors (2020).

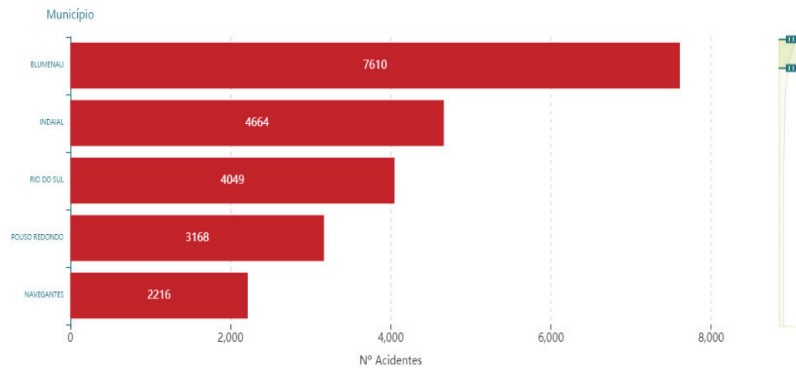


Figure 10. Accidents per municipality on BR 470 in Santa Catarina
 Source: Prepared by the authors (2020).

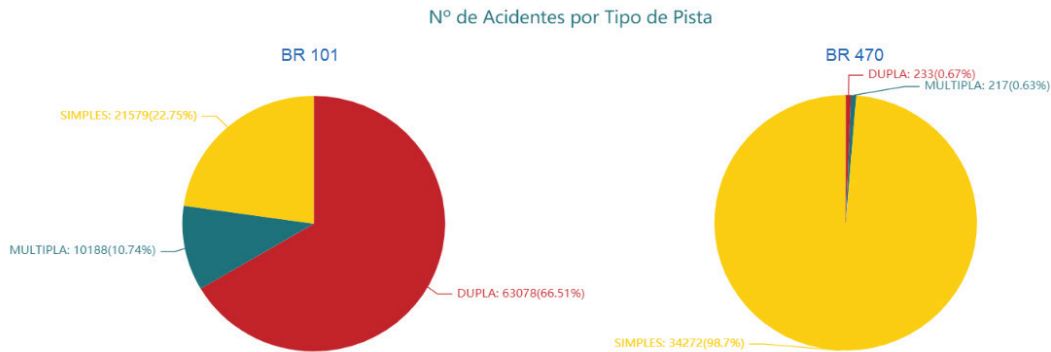


Figure 11. Type of lane on which accidents happen on BR 470 and BR 101 in Santa Catarina
 Source: Prepared by the authors (2020).

tims (BR 101, with 44.11%, and BR 470, with 45.96%), and third are the fatal victims (BR 101, with 2.17%, and BR 470, with 3.34%).

One of the hypotheses we can think of for this number of accidents without victims is the public's policies for the safety and preservation of life through traffic awareness policies. Among the measures, it is worth highlighting Resolution No. 277/2008 of the National Traffic Council (CONTRAN, 2008), known as the "car seat law," and Law No. 11.705/2008, known as the "dry law," which came into force in 2008; Law No. 12.760/2012, which instituted an increase in the fine since the dry law was not effective, i.e., did not reduce accidents involving drunkenness (Brazil, 2012); Law No. 12.971/2014, which increases the punishment regarding improper overtaking and behavior; and Law No. 13.821/2016, in which the values of fines for all traffic infractions were modified, having an increase of more than 50% (Brazil, 2014; 2016).

(v) What are the main causes of accidents on a given highway?

To identify the causes of the most frequent accidents, the procedure is the same as the one performed so far; how-

ever, the characteristic to be selected is "conditioning," and the variable chosen is "accident cause." Then, in the new query, the variable must be changed to "accident type." We performed this sequence for both BR 101 and BR 470.

When analyzing the BR 101, we can see in **Figure 13** that the four biggest causes of accidents are: lack of attention to driving (43,417 accidents), other (14,511 accidents), not keeping a safe distance (10,061 accidents), and incompatible speed (5,879 accidents). Regarding the type of accident (**Figure 14**), it is worth noting that in first place are rear-end collisions (29,527 accidents), followed by side collisions (16,470 accidents).

Following the analysis of BR 470, we can see in **Figure 15** that the four main causes of accidents are: inattention while driving (15,086 accidents), others (3,696 accidents), not keeping a safe distance (3,584 accidents), and incompatible speed (3,419 accidents). Regarding the type of accident (**Figure 16**), it is worth noting that in the first place are rear-end collisions (8,988 accidents), followed by transverse collisions (5,587 accidents), and then side impact collisions (5,177 accidents).

On both highways, inattention while driving is the main cause of road accidents. We can correlate it to the largest type of accident: rear-end collisions. Studies by the Pan American Health Organization (PAHO, 2019, n.p.) reveal that the “distraction caused by cell phones is a growing concern for traffic safety.” Another aspect is the increase in the number of accidents. Drivers who use a cell phone behind the wheel “are about four times more likely to be involved in a crash” (PAHO, 2019, n.p.).

(vi) What is the most common age range of people involved in accidents in a given state?

To finalize our analysis and show the dashboard’s potential for public policies, we present the graph regarding the age range of those involved in the accidents. The search procedures are the same as previously applied; we only changed the characteristic of the accidents to “involved” and the variable selected was “people’s age” (Figures 17 and 18).

We can observe in BR 101 (Figure 17) and BR 470 (Figure 18) that the age range of those involved is between 25 and 30 years old, corresponding to 28,270 and 10,451 accidents, respectively. According to Albertini *et al.* (2020, p. 2,800),

land transport accidents in Brazil constitute a public health problem and are “the leading cause of death of young adults between 20 and 29 years old, especially males.” The authors recognized a pattern of epidemiological characteristics and the profile of the traffic accident victims: “According to DATASUS, 82.38% of the victims of fatal accidents are men, and 17.62% are women; most of these victims are young adults, aged between 20 and 49 years, and owners of motorcycles” (Albertini *et al.*, 2020, p. 2,802).

In this section, we have presented some of the dashboard’s functionalities for accessing public data related to the population’s daily life, as they will reflect on public policies, from traffic awareness and education to the implementation of improved access to the BRs and signaling, among other aspects.

CONCLUSION

The Access to Information Law of 2011 was important to oblige public entities and private non-profit organizations to publish their data. The law covers federal, state, district, and municipal agencies and entities of the executive, legislative,

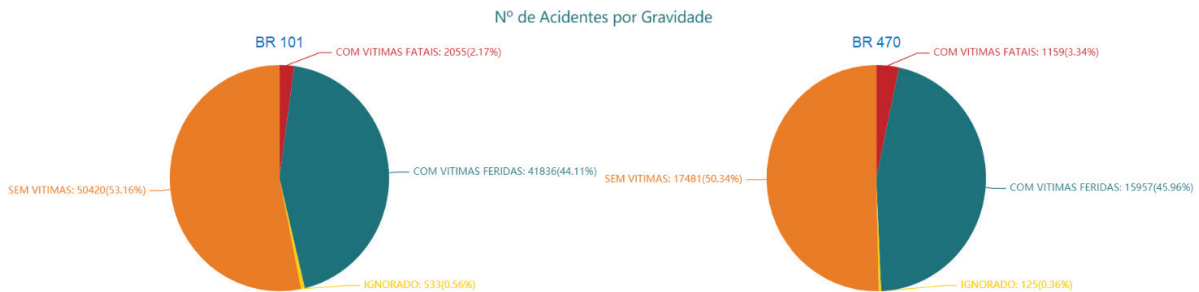


Figure 12. Severity of accidents in Santa Catarina on stretches of BR 470 and BR 101
 Source: Prepared by the authors (2020).

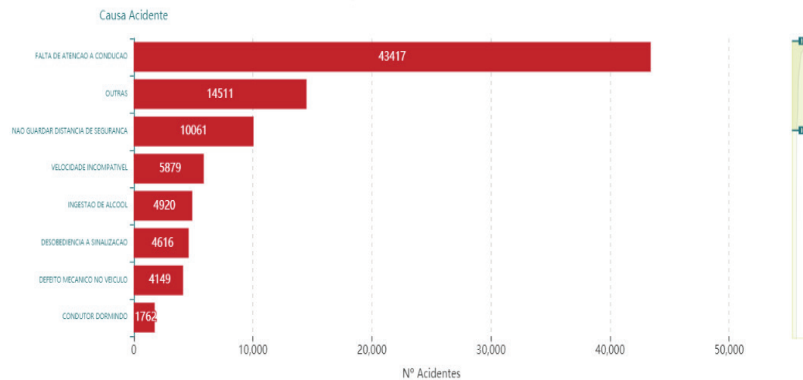


Figure 13. Causes of accidents in Santa Catarina on stretches of BR 101
 Source: Prepared by the authors (2020).

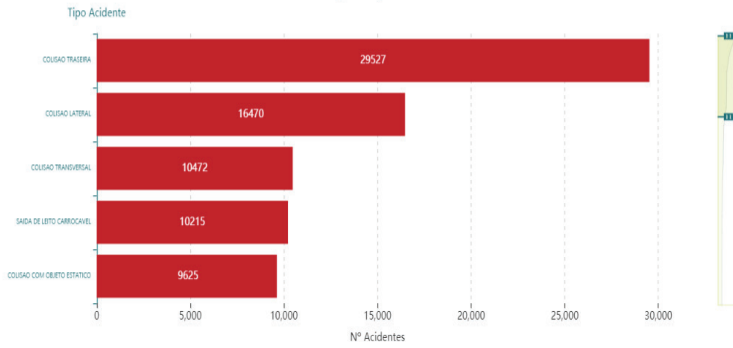


Figure 14. Types of accidents in Santa Catarina on stretches of BR 101
 Source: Prepared by the authors (2020).

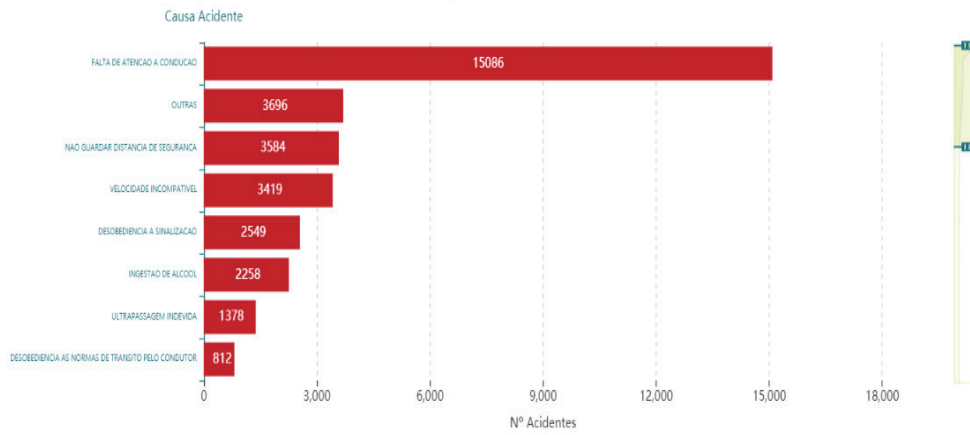


Figure 15. Causes of accidents in Santa Catarina on stretches of BR 470
 Source: Prepared by the authors (2020)

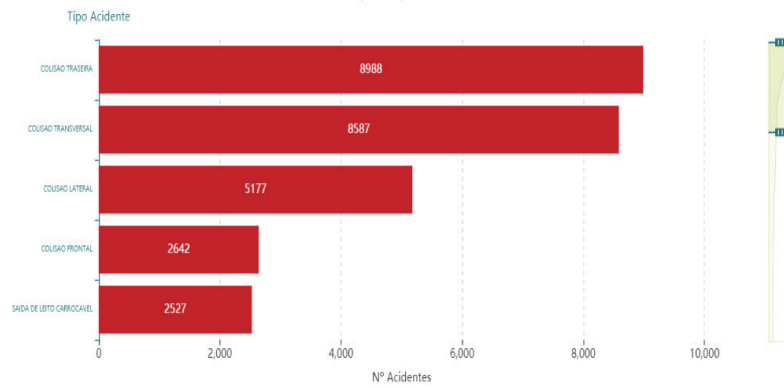


Figure 16. Types of accidents in Santa Catarina on stretches of BR 470
 Source: Prepared by the authors (2020)

and judicial branches. One of these data is related to accidents on federal highways in Brazil, available on the DPRF website, which was the subject of analysis in this work.

It was evaluated how the data on the DPRF website are made available and if they follow the eight principles suggested by the Brazilian Open Data Portal (PBDA, 2020), namely: complete, primary, current, accessible, machine processable, non-discriminatory access, non-proprietary format, and free of licenses. It was found that they partially follow the principles, mainly because they do not meet the non-proprietary format (the files are compressed in .rar format) and because the data is updated every three months. These observations should be communicated to the organization's ICT area for better adapting to the principles that govern open data. Regarding the seven additional principles suggested by the Open Government Data (OGD, 2007): online and free; permanent; reliable; openness presumption; documented; safe to open; and designed with public contribution, it was found that the DPRF website meets five of these, with the principles "Reliable" and "Designed with the public contribution" not being met.

Then, a data science design methodology was used to build an informative dashboard using R software (open source). This dashboard aims to improve access to information about accidents on federal highways in Brazil.

In the example of dashboard usage by the users, the state of SC and the BRs 101 and 470 were selected. As a result, one can see that the accidents concentrate in the metropolitan regions. Thus, the following questions arise: Is there a lack of signaling on the federal highway on these routes? Is it the several access points to the city? Is it the intense flow of vehicles? These questions based on the dashboard can contribute to the different spheres of public power, aiming at the improvement and safety of the population. Thus, the questions surrounding this theme are numerous, but using open public data to improve the social welfare of society is a right.

For future work, we suggest an accident prediction model based on the open data made available by the DPRF. These models can contribute to public administration policies because, from the metrics and general statistics, it will be

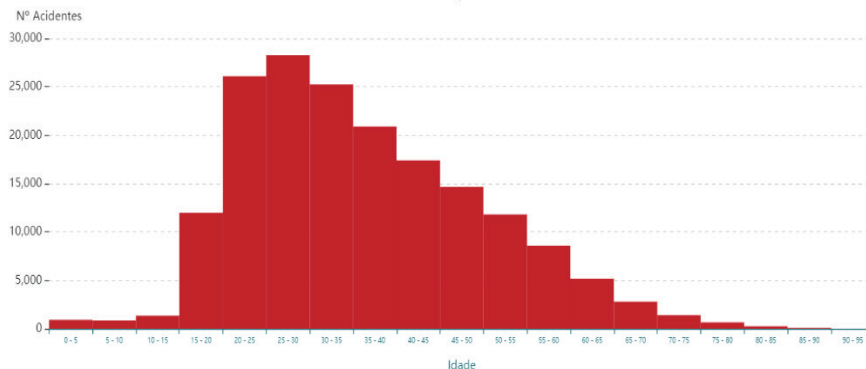


Figure 17. Age of people involved in accidents on BR 101 in Santa Catarina
Source: Prepared by the authors (2020)

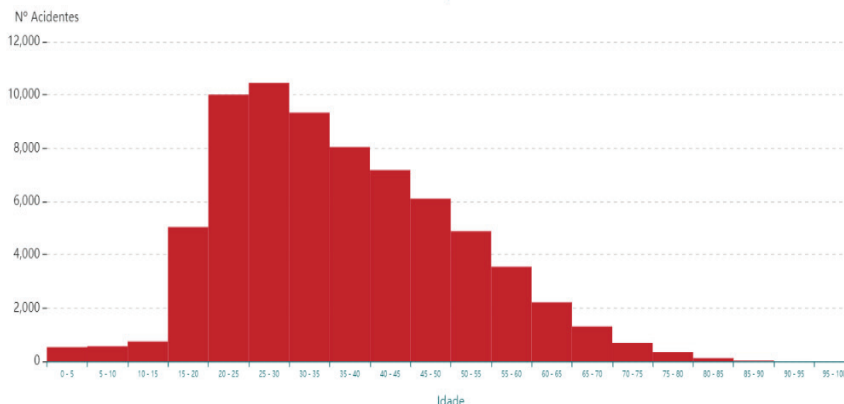


Figure 18. Age of people involved in accidents on BR 470 in Santa Catarina
Source: Prepared by the authors (2020)

possible to verify the focus of greater vulnerability, i.e., the causes and places of higher accident concentration.

REFERENCES

- Albano, C.S., Leaes Júnior, W. & Corso, K.B. (2019), “Dados abertos: um estudo em trabalhos acadêmicos no contexto brasileiro”, *Navus*, vol. 9, no. 2, pp. 135-47.
- Albertini, A.B.N., Garcia, T.V., Paulo, L.G., Toledo, E.R.S., Charlo, P.B. & Silva, M. (2020), “Caracterização epidemiológica e sociodemográfica de acidentes de trânsito: uma revisão integrativa da literatura”, *Saúde Coletiva (Barueri)*, vol. 10, no. 55, pp. 2797-814.
- Allaire, J.J., Xie, Y., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., Wickham, H., Cheng, J., Chang, W. & Iannone, R. (2020), *Rmarkdown: Dynamic Documents for R*, version 2.1, R-Packages, disponível em: <https://cran.r-project.org/web/packages/rmarkdown/index.html> (acesso em 10 jan. 2020).
- Attard, J., Orlandi, F., Scerri, S. & Auer, S. (2015), “A systematic review of open government data initiatives”, *Government Information Quarterly*, vol. 32, no. 4, pp. 399-418.
- Boulton, G. (2014), “The open data imperative”, *Insights*, vol. 27, no. 2, pp. 133-8.
- Boulton, G., Rawlins, M., Vallance, P. & Walport, M. (2011), “Science as a public enterprise: The case for open data”, *The Lancet*, vol. 377, no. 9778, pp. 1633-5.
- Brasil (2011), Lei Ordinária n. 12.527, de 18 de novembro de 2011, *Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei n. 11.111, de 5 de maio de 2005, e dispositivos da Lei n. 8.159, de 8 de janeiro de 1991; e dá outras providências*, Presidência da República, Brasília, DF.
- Brasil (2012), Lei nº 12.760, de 20 de dezembro de 2012, *altera a Lei nº 9.503, de 23 de setembro de 1997, que institui o Código de Trânsito Brasileiro*, Presidência da República, Brasília, DF.
- Brasil (2014), Lei nº 12.971, de 9 de maio de 2014, *Altera os artigos 173, 174, 175, 191, 202, 203, 292, 302, 303, 306 e 308 da Lei nº 9.503, de 23 de setembro de 1997, que institui o Código de Trânsito Brasileiro, para dispor sobre sanções administrativas e crimes de trânsito*. Presidência da República, Brasília, DF.
- Brasil (2016), Lei nº 13.281, de 4 de maio de 2016, *Altera a Lei nº 9.503, de 23 de setembro de 1997 (Código de Trânsito Brasileiro), e a Lei nº 13.146, de 6 de julho de 2015*. Presidência da República, Brasília, DF.
- Braunschweig, K., Eberius, J., Thiele, M., Lehner, W. & Dresden, T.U. (2012), “The state of open data. limits of current open data platforms”, artigo apresentado no WWWC 2012: World Wide Web Conference, Lyon, France, 21 de abr. 2012.
- Chang, W. & Ribeiro, B.B. (2018), “Shinydashboard: Create Dashboards with ‘Shiny’”. R package version 0.7.1.”, disponível em: <https://CRAN.R-project.org/package=shinydashboard> (acesso em: jun. 2020).
- Chang, W., Cheng, Allaire, J.J., Sievert, C., Schloerke, B., Xie, Y., Allen, J., McPherson, J., Dipert, A., Borges, B. (2020), *shiny: Web Application Framework for R*, version 1.4.0.2., R-Packages, disponível em: <https://cran.r-project.org/web/packages/shiny/index.html> (acesso em: 13 jun. 2019).
- Cooper, N. (2017), *Reader: Suite of Functions to Flexibly Read Data from Files*, version 1.0.6.2, R-Packages, disponível em: <https://cran.r-project.org/web/packages/reader/index.html> (acesso em: 10 dez. 2019).
- Dancho, M. & Vaughan, D. (2020), *sweep: Tidy Tools for Forecasting*, disponível em: <https://cran.r-project.org/web/packages/sweep/index.html> (acesso em: 10 dez. 2019).
- Departamento de Polícia Rodoviária Federal (2019a), *Dados Abertos*, disponível em: <https://www.pr.f.gov.br/portal/dados-abertos> (acesso em: 06 dez. 2019).
- Departamento de Polícia Rodoviária Federal (2019b), *Dicionário de Variáveis – por ocorrência*, disponível em: <https://www.pr.f.gov.br/portal/dados-abertos/acidentes/icionario-de-variaveis-por-ocorrencia> (acesso em: 06 dez. 2019).
- Departamento Nacional de Infraestrutura de Transportes (2020), *Nomenclatura das Rodovias Federais*, disponível em <https://www.gov.br/dnit/pt-br/rodovias/rodovias-federais/nomenclatura-das-rodovias-federais> (acesso em: 10 nov. 2020).
- Diniz, V. (2010), “Como conseguir dados governamentais abertos”, artigo apresentado no III Congresso CONSAD de Gestão Pública, Brasília, DF, 15-17 mar. 2010.
- Gil, A.C. (2017), *Como elaborar projetos de pesquisa*, 4a ed., Atlas, São Paulo.
- Gorges, L. (2019), *Verão Santa Catarina 2019/2020 mobiliza 9,2 mil profissionais da segurança e prevê investimento de R\$ 41,1 milhões*, Agência Catarinense de Notícias, Florianópolis, SC, disponível em: <https://www.sc.gov.br/noticias/temas/verao-santa-catarina/verao-santa-catarina-2019-2020-mobiliza-9-2-mil-profissionais-da-seguranca-e-preve-investimento-de-41-1-milhoes> (acesso em: 30 out. 2020).
- Governo do Brasil (2020), *Brasil registra queda em número de mortes no trânsito*, Governo do Brasil, Brasília, DF, disponível em <https://www.gov.br/pt-br/noticias/transito-e-transportes/2020/09/brasil-registra-queda-em-numero-de-mortes-no-transito> (acesso em: 05 nov. 2020).
- Harrison, T. M., Harrison, T.M., Guerrero, S., Burke, G.B., Cook, M.E., Cresswell, A.M., Helbig, N., Hrdinová, J., & Pardo, T.A. (2012), “Open government and e-government: democratic

challenges from a public value perspective”, *Information Polity*, vol. 17, no. 2, pp. 83-97.

Henry, L. & Wickham, H. (2020), *purrr: Functional Programming Tools*, version 0.3.3., R-Packages, disponível em: <https://cran.r-project.org/web/packages/purrr/index.html> (acesso em: 03 mai. 2020).

Instituto de Pesquisa Econômica Aplicada (2015), *Acidentes de Trânsito nas rodovias federais brasileiras: caracterização, tendências e custos para a sociedade*, Ipea, Brasília, disponível em: http://www.ipea.gov.br/portal/index.php?option=com_content&view=article&id=26277 (acesso em: 15 jun. 2020).

Open Government Data (2007), *The 8 Principles of Open Government Data*, Sebastopol, California, disponível em: <https://opengovdata.org/> (acesso em: 17 jul. 2020).

Open Government Partnership (2011), *Declaração de Governo Aberto*, disponível em: <http://www.governoaberto.cgu.gov.br/central-de-conteudo/documentos/arquivos/declaracao-governo-aberto.pdf> (acesso em: 16 jul. 2020).

Open Knowledge Foundation (2020), *What is Open Data?*, Open Knowledge Foundation, disponível em: <http://opendatahandbook.org/guide/en/what-is-open-data/> (acesso em 15 jul. 2020).

Organização Pan-Americana da Saúde (2019), *Folha informativa - Acidentes de trânsito*, Organização Pan-Americana da Saúde, disponível em: https://www.paho.org/bra/index.php?option=com_content&view=article&id=5147:acidentes-de-transito-folha-informativa&Itemid=779 (acesso em: 01 out. 2020).

Peña, C.C., Tani, V., Kaesemodel, L., Otto, G.G., Wutke, J.D. & Andrade, R.F. (2008), *Identificação dos segmentos críticos das Rodovias Federais de Santa Catarina*, Departamento Nacional de Infraestrutura de Transportes, Universidade Federal de Santa Catarina, Florianópolis, SC.

Pinho, M.D.C. & Silva, M.P.A. (2019), “Governo aberto e dados abertos governamentais: um mapeamento e sistematização da produção acadêmica”, *Comunicação & Inovação*, vol. 20, no. 43, pp. 3-25.

Portal Brasileiro de Dados Abertos (2020), *O que são dados abertos?*, Portal Brasileiro de Dados Abertos, disponível em: <http://dados.gov.br/pagina/dados-abertos> (acesso em: 29 mai. 2020).

Possamai, C.A.P. (2017), Estudos de pontos críticos de acidentes de trânsito nas Rodovias Federais de Santa Catarina e melhorias viárias de baixo custo, Monografia de Especialização em Perícia de Acidentes de Trânsito, Instituto Federal de Santa Catarina, Florianópolis, SC.

R Core Team (2020), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, disponível em: <http://www.R-project.org/> (acesso em: 03 jan. 2020).

Robinson, D., Hayes, A., Couch, S., Psit Software, Patil, I., Chiu, D., Gomez, M., Demeshev, B., Menne, D., Nutter, B., Johnston, L., Bolker, B., Briatte, F., Arnold, J., Gabry, J., Selzer, L., Simpson, G., Preusser, J., Hesselberth, J. & Wickham, H. (2020), *broom: Convert Statistical Analysis Objects into Tidy Tibbles*, version 0.5.5., R-Packages, disponível em: <https://cran.r-project.org/web/packages/broom/index.html> (acesso em: 05 jan. 2020).

Sievert, C., Parmer, C., Hocking, T., Chamberlain, S., Ram, K., Corvellec, M., Despouy, P., Bruggemam, S., Inc, P.T. (2020), *plotly: Create Interactive Web Graphics via 'plotly.js'*, version 4.9.2.1, R-Packages, disponível em: <https://cran.r-project.org/web/packages/plotly/index.html> (acesso em: 02 abr. 2020).

Silva, C.F., Santos, E.M.F., Chaves, M.C., Vaz, W. & Balaniuk, R. (2014), “Dados abertos: uma estratégia para o aumento da transparência e modernização da gestão pública”, *Revista TCU*, vol. 131, pp. 22-9.

Tischer, V. (2019), “O custo social e econômico dos acidentes de trânsito com pedestres e ciclistas: estudo de caso do estado de Santa Catarina, Brasil”, *Revista Brasileira de Gestão Urbana*, vol. 11, pp. 1-14.

Tribunal de Contas da União (2015), *Cinco motivos para a abertura de dados na administração pública?*, TCU, Brasília, disponível em: <https://portal.tcu.gov.br/biblioteca-digital/cinco-motivos-para-a-abertura-de-dados-na-administracao-publica.htm> (acesso em: 17 jul. 2020).

Veljković, N., Bogdanović-Dinić, S. & Stoimenov, L. (2014), “Benchmarking open government: An open data perspective”, *Government Information Quarterly*, vol. 31, no. 2, pp. 278-90.

Vieira, R. (2020), *A rodovia do Vale: Trevo do Badenfurt é novo desafio da duplicação da BR-470 em SC*, disponível em: <https://ndmais.com.br/economia-sc/a-rodovia-do-vale-trevo-do-badenfurt-e-novo-desafio-da-duplicacao-da-br-470-em-sc/> (acesso em: 02 dez. 2020).

Wickham, H. & Bryan, J. (2019), *Readxl: Read Excel Files*, version 1.3.1, R-Packages, disponível em: <https://cran.r-project.org/web/packages/readxl/index.html> (acesso em 12 nov. 2020).

Wickham, H. & Grolemund, G. (2017), *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*, O'Reilly Media, Gravenstein Highway North, Sebastopol, CA.

Wickham, H. & Henry, L. (2020), *tidyr: Tidy Messy Data*, version 1.0.2, R-Packages, disponível em: <https://cran.r-project.org/web/packages/tidyr/index.html> (acesso em: 19 fev. 2020).

Wickham, H. & Müller, K. (2019), *DBI: R Database Interface*, version 1.1.0, R-Packages, disponível em: <https://cran.r-project.org/web/packages/DBI/index.html> (acesso em: 18 nov. 2019).

Wickham, H. (2020), *hhtrs: Tools for Working with URLs and HTTP*, version 1.4.1, R-Packages, disponível em: <https://cran.r-project.org/web/packages/hhtr/index.html> (acesso em: 13 nov. 2019).

Wickham, H., Chang, W., Henry, L., Pedersen, T.L., Takahashi, K., Wilke, C. Woo, K., Yutani, H. & Dunnington, D. (2020a), *ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*, version 3.3.0, disponível em: <https://cran.r-project.org/web/packages/ggplot2/index.html> (acesso em: 25 mar. 2020).

Wickham, H., François, R., Henry, L., Müller, K. & Vaughan, D. (2020b), *dplyr: A Grammar of Data Manipulation*, version 0.8.5, disponível em: <https://cran.r-project.org/web/packages/dplyr/index.html> (acesso em: 20 mar. 2020).

World Health Organization (2018). *Global status report on road safety 2018*, World Health Organization, Genebra.

Received: February 22, 2021

Approved: June 5, 2023

DOI: 10.20985/1980-5160.2023.v18n1.1706

How to cite: Miranda, R.G., Konrath, A.C., Zatelli, M.R., Hoffmann, Y.T., Silva, S.A. (2023). A dashboard for decision-making considering open data from federal highways. *Revista S&G* 18, 1. <https://revistasg.emnuvens.com.br/sg/article/view/1706>