HealthDashboard: A Urban Public Health Geospatial Visualization Platform

Gabriely Pereira* Lucas Stankus† Gustavo Carlos‡ Eduardo Pinheiro§ Rafael Manzo¶
Kelly Braghetto‖ Fabio Kon** Paulo Meirelles††

Free/Open Source Software Competence Center
Institute of Mathematics and Statistics (IME)
University of São Paulo (USP)

Abstract

Public healthcare systems generate large amounts of heterogeneous data that can provide valuable insights to inform public policy design. However, extracting relevant information from extensive heterogeneous datasets might be challenging. To address this problem, in a government-academia collaboration, we are developing an interactive visual dashboard for large-scale data analysis based on the Brazilian National Health System (SUS) hospitalization data. Its software architecture enables integration with the Hospital Information System (SIH-SUS) datasets from any region of Brazil so that health professionals can use it in hundreds of different cities. We defined an architecture that tames code complexity and brings modularity to the system. The platform processes SIH-SUS data and stores it into a geolocated relational database. Expert users can then perform advanced queries on the data with composite filters. Results are then displayed via multiple map visualizations, graphs, and tables. We expect that this open-source platform will become a useful tool for science-based public health policymaking, influencing Brazilian public managers in the future to adopt an evidence-based, data-driven approach to healthcare management.

Index Terms: Human-centered computing—Visualization; Software and its engineering—Open source model

1 Introduction

With the increasing urbanization, mainly in developing countries, urban resources' efficient management has become a crucial point for modern cities' quality of life. In this scenario, recent advances in computing technologies can provide practical tools to support urban information analysis and contribute to evidence-based public policymaking. This use of technology is what defines *Smart Cities*.

Smart cities aim to improve urban efficiency and quality of life, contributing to city sustainability, mobility, healthcare, and accessibility [6, 10]. Concerning healthcare, there is a significant growth in the generated and stored data from health facilities [16]. A large body of data can provide relevant information that can be used by public managers to support the design of public policies [12]. However, analyzing extensive collections of heterogeneous data to develop useful insights is often a difficult task. A robust approach to achieve that is to employ carefully designed visualizations that display large amounts of data to expert users, enabling them to have useful insights. Our research group has collaborated with Brazilian public health officials to develop a large-scale data visualization platform. The platform allows for the analysis of urban health data via modern techniques for data visualization. It uses a broad set of hospitalization and health facilities data, the municipality geographic structure, and sociodemographic information.

Implementing Smart Cities initiatives in a developing country such as Brazil is a challenging endeavor. Since the resources are scarce and local governments have a minimal budget, there is a clear need for a more scientific approach to policymaking. Although the local staff is aware of this need, convincing higher-level managers and politicians that public management changes towards a more scientific approach are urgent is an arduous task. We expect our open-source platform to become a useful tool for evidence-based public health policymaking, influencing public managers in the future to adopt a data-driven approach to management.

The platform produces visualizations from different perspectives of the Hospital Information System (SIH) dataset provided by the Brazilian National Health System (SUS). This dataset contains information such as diagnosis, hospital location, geo-anonymized patient location, and hospitalization data, among other data. The platform processes multiple georeferenced data allowing their visualization on a dynamic, interactive map, enabling identifying patterns in geographic and temporal variations.

The platform is capable of dealing with any SIH-SUS dataset. It processes CSV data obtained from government systems and stores SIH-SUS data into a relational database. The data is then available for authenticated users through a user-friendly interface that requires no SQL knowledge. A healthcare expert user performs advanced queries applying composite filters over the attributes stored in the database. Our implementation uses caching techniques to provide timely responses while working with large datasets. The system displays query results in tables, charts, and multiple map visualizations, such as clustering and heat maps. Resulting data can also be exported into PDF or CSV files that health professionals can use to create reports and support decision making.

In the paper's remainder, we discuss the related work, introduce the SIH-SUS dataset, present the visualization platform's requirements, architecture, and features, and make the concluding remarks.

*e-mail: gabriely@ime.usp.br
†e-mail: lucas.p.stankus@usp.br
‡e-mail: gustavo.10.dmc@usp.br
§e-mail: edup@ime.usp.br
¶e-mail: manzo@ime.usp.br
‖e-mail: kellyrb@ime.usp.br
**e-mail: kon@ime.usp.br
††e-mail: paulormm@ime.usp.br

2 Related Work

Several works have already addressed the visualization of Brazilian public health data from SUS systems. However, most of them relate to the use of Geographic Information Systems (GIS) to investigate the epidemiology of a specific disease, as made by Hertz et al. [3] and Marinho et al. [8]. More general initiatives are the PCDaS platform and the Conecta SUS project.

The PCDaS – *Plataforma de Ciência de Dados aplicada à Saúde* (Data Science Platform Applied to Health)[1] was created by the Institute of Communication and Scientific and Technological Information in Health (Icict) of the Oswaldo Cruz Foundation (Fiocruz), in partnership with the National Laboratory for Scientific Computing (LNCC), in 2016. The platform offers open services for storage, management, analysis, and dissemination of public health data and gathers a large SUS data volume. PCDaS has dashboards with basic capabilities for visual analysis but does not use maps to their full potential to support geospatial data exploration. The dashboards are implemented in Kibana, an open-source tool for visualizing data stored in Elasticsearch clusters.

The Conecta SUS [2] project was developed by the Zilda Arns Neumann Center for Strategic Information and Decision-Making in Health from the State Health Secretariat of Goiás and is being replicated in other states in the country. The Conecta SUS integrates data from multiple systems to subsidize the Secretariat's coordination of actions and policies. It uses data warehousing and business intelligence tools to provide real-time visual information and historical analysis. Its infrastructure is based on a PostgreSQL/PostGIS database, an ETL process implemented in Pentaho Data Integration, OLAP cubes in Mondrian, and monitoring panels designed in Pentaho Community.

Differently from these solutions, the platform we present in this paper focuses on the geospatial presentation of hospitalization data. It uses different types of dynamic, interactive maps to provide rich visualizations of georeferenced data from the Hospital Information System (SIH-SUS). These visualizations show healthcare resource usage and help identify health service demands while supporting epidemiologic studies.

Moreover, several initiatives were created in healthcare administration, such as ours, which developed tools for assisting the healthcare decision-making process. The following paragraphs run through examples of such initiatives related to our platform.

DataScope [5] is a visual querying tool for exploring large datasets. It provides multiple visual dashboards with querying and filtering options, allowing health researchers to cut down data to create relevant subsets or cohorts for their research. The resulting visualizations can be saved to be incorporated in reports or for continuous analysis. Another initiative similar to DataScope is the prototype done during the design study headed by Zhao [19], in which the developers worked with health researchers' feedback to design a querying solution with intuitive user interfaces that were less mentally stressful to use.

PandemCap [17] is a visual analytics support tool for pandemic management. It employs epidemic modeling in conjunction with data visualization to improve the public healthcare decision-making process during an outbreak. It allows government officials to compare multiple scenarios according to which interventions are made. Within these scenarios, they have a vast arrange of information regarding the disease spread, *e.g.* the number of hospital beds available by health facility and region. This information enhances understanding of the situation and its possible outcomes, allowing better-informed decisions.

The Patient-Provider Geographic Map [18] is a geospatial visualization tool targeted to health administrators. The tool provides filtering options based on health care providers and diagnostic groups and displays selected patient's data through ZIP-code level choropleth maps. This visualization allows findings over underserved or overserved populations and patient's health care provider preferences.

The Community Health Map [13] is a web application designed to analyze the United States' public health. Its differential from similar tools is the ability to work simultaneously with healthcare and demographics datasets, which allows findings that would not be possible in other tools. Users can filter the expanded dataset of Community Health Map by parameters not related to healthcare but that significantly impact it. The availability to the public, self-reported overall quality of health services, and the county's poverty rate are examples of filters that can be useful for government health officials.

3 Dataset

Founded in 1991, the SUS Information Technology Department (DATASUS) of the Brazilian Ministry of Health is responsible for providing information systems and IT support for SUS' agencies. Since its creation, DATASUS developed more than 200 different software systems to grow and strengthen SUS. Along with software development, the department also grants digital infrastructure for supported agencies and is bound to collect and store data generated by SUS' facilities across all Brazilian regions. The collected health data is free to the general public and is widely available through the DATASUS website[2].

One of the oldest running systems in the department is the SUS Hospital Information System (SIH-SUS), the first DATASUS system to have data capture implemented in microcomputers and decentralized to its users. Data entries are obtained by many filled forms from the country's health facilities called Hospital Admission Authorization (AIH). An AIH, if correctly filled, should log all relevant information about the health procedure, including, but not limited to, administrative information (*e.g.*, the Regional Health Coordination responsible for the area where the procedure took place), data about the patient, the health facility where the patient was admitted, and health logs about the procedure itself (*e.g.*, patient diagnosis and the procedure's specialty).

The SIH-SUS dataset is provided via TabNet[3] by the Federal Brazilian government. TabNet is a tabulation and georeferencing system for public health data. It processes the tabulated data published by the Ministry of Health, focusing on facilitating the calculation of health indicators.

A submitted AIH form is processed and, by design, becomes available on the SIH-SUS dataset within one month after it has been filled in. The SIH-SUS keeps data on about 70% of all country's hospitalizations. With this high percentage of cases covered, the morbidity and mortality rates converge closely to their real values, so it is essential to assess the SUS healthcare assistance's importance and plan precautionary measures. These data provide to public health managers information that supports decision-making. For example, SIH-SUS data may show regions lacking public hospital

[1] https://bigdata.icict.fiocruz.br/

[2] https://datasus.saude.gov.br

[3] www2.datasus.gov.br/DATASUS/index.php?area=02

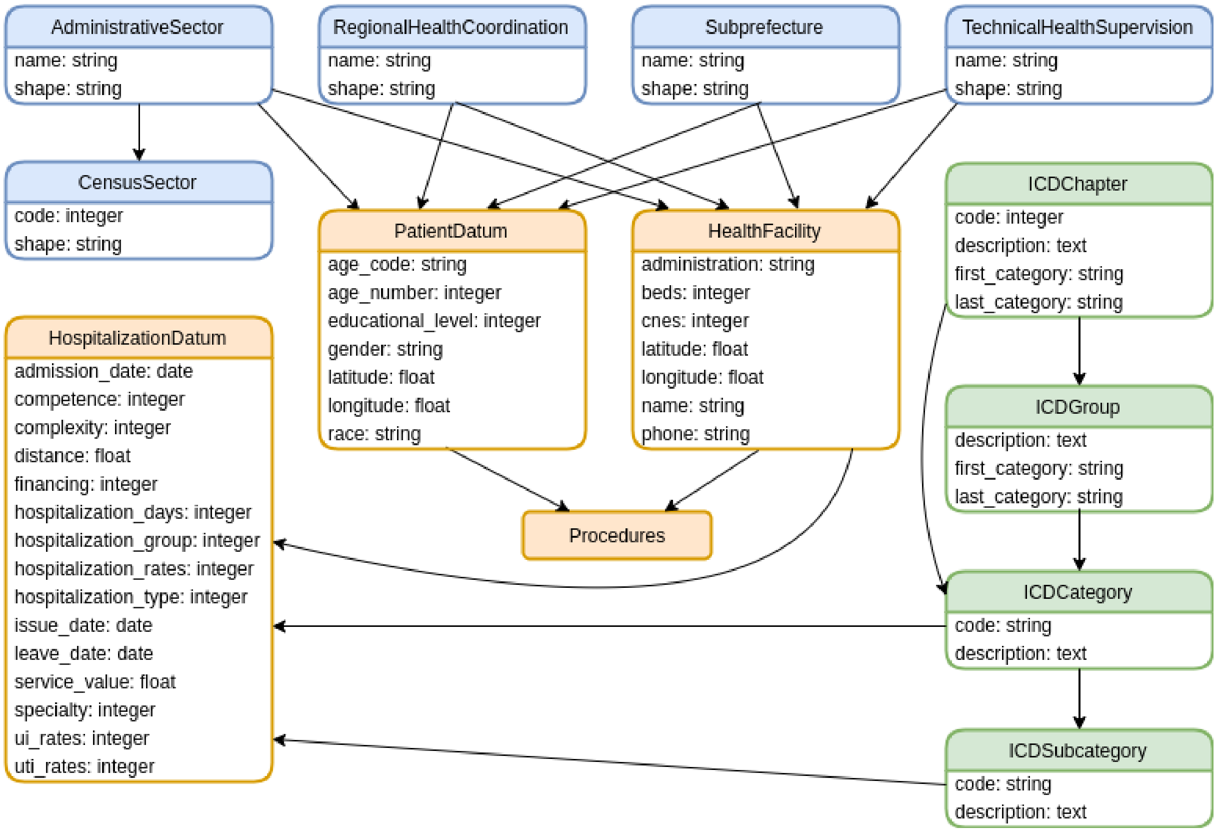

Figure 1: Models of the platform and their relationships

beds or services in specific medical specialties.

As stated before, the SIH-SUS dataset is composed of data from all over Brazil, collected in a decentralized manner by each Health Secretariat, and sent to the Brazilian Health Ministry. However, each secretariat can extract the data from the SIH-SUS dataset and complement it with some information to favor further analysis, and this change can make the structure of the local dataset subtly vary from one region to another. For example, our Health Secretariat partner pre-processed the SIH-SUS dataset removing the corrupted data, anonymizing the entire dataset, and inserting the geolocation (in terms of latitude and longitude) of the census tract centroid of the patient location (*i.e.*, geo-anonymizing the patient location).

Since 1995, the SIH-SUS datasets retrieved more than 285 million hospitalization records according to the DATA-SUS TabNet system. Our partner's pre-processed SIH-SUS dataset covers 554,202 hospitalization records of a city from January 2015 to December 2015. This large dataset was used to conduct several tests on the visualization platform presented in this paper.

We designed the platform architecture based on the pre-processed SIH-SUS dataset, with two main models: *Patient Data* and *Hospitalization Data*. *Patient Data* model contains the patient anonymized data (age, gender, educational level, and others) and location, originating from the *Location* model, which stores the name and the geographic data structure of the municipality subdivisions, such as subprefectures, administrative sectors, and census tracts. *Hospitalization Data* model is composed of the hospitalization data (admission date, discharge date, hospitalization complexity, and others); the hospitalization diagnosis, represented by the *Diagnosis* model, which is responsible for cataloging the ICD-10 (International Classification of Diseases) categories, subcategories, groups, and chapters; the health facility data (number of beds, phone number, and others) obtained from the National Database on Health Units (SCNES); and the health facility location, represented by the *Location* model. The complete model structure of the platform is represented in Figure 1.

4 Development Approach

Our platform focuses on promoting better analysis for georeferenced hospitalization data through a data visualization dashboard. Its implementation aims to deal with any SIH-SUS dataset, making possible the analysis of public hospitalizations from any region of Brazil.

A useful data visualization dashboard should present essential information, concisely summarize all the data that matters, and be interactive, allowing easy data exploration. The platform should also enable data import and export in different formats, such as PDF and CSV. Besides, for the proper analysis of georeferenced datasets, the dashboard should provide features for the spatial visualization of data. Our platform meets all these requirements while supporting the analysis of the SIH-SUS large dataset.

We developed the platform in a government-academia collaboration. This scenario is often challenging, and both parties need to strive to increase the chances of success. Interinstitutional conflicts should be mitigated to obtain a successful collaboration [15]. In our collaboration with a Brazilian Health Secretariat, we adopted practices from agile methods and Open Source Software (OSS) communities, following the lessons learned by Wen et al. [15] to involve the Brazilian public health officials (expert users) in the development process. They participated, updated, and created issues (features and bugs) into the project repository based on the requirements defined by them. The government officials and expert users interacted directly with the development team via the project's repository and received on-the-fly notifications of advances and difficulties we were facing while receiving frequent feedback from their homologation tests. We also had monthly strategy meetings with them, which helped us identify the best technical approach to fit their needs and the available resources.

Development activities occurred based on project maintainers and developers' contributions. Project maintainers can update the main source code, while project developers can update the code through merge requests (MR). In this way, we adopt a development workflow based on OSS communities, intending to provide the source code quality through submitted code inspections [11]. In summary, each contributor can make a copy (fork) of the project repository, change the source code, and return these changes to the principal repository via merge requests. From that point, a maintainer should review the submitted code and, eventually, request modifications and improvements. Finally, when the submitted code is within the project's quality standards, a maintainer can accept the MR and incorporate the code changes. Our project's code and issues are available at a GitLab repository [4] under the Mozilla Public License 2.0.

5 Architecture

We adhered to well-established design principles to build a feasible, robust, and concrete dashboard application architecture. These principles contribute to the software extensibility and maintainability. The platform requires `modularity` since it facilitates the code's comprehension and increases its cohesion due to the division in logical components [14]. Modularity also improves communication between parts of the code and facilitates full test coverage because it is easier to test small modules than the entire code base. The `reuse of OSS projects` can improve faster code development, save costs, and enhance code reliability [11]. However, it is essential to be aware of the package quality and to use well known free projects with an active developer community. We also follow the *`Don't Repeat Yourself`* `principle` (DRY) in our development. It states that every piece of knowledge needs to have a single representation [4]. Furthermore, a piece of knowledge can be either the build system, the database schema, the tests, or even the documentation.

Based on these principles, we adopted the MVC (*Model-View-Controller*) architectural style [7] to implement our platform, considering system maintainability and extensibility as fundamental concerns. `Model` components are those parts of the system application that simulate the application domain. In our case, models represent the SIH-SUS dataset inside the platform. The hospitalization *Procedure* model is the main model in the platform. `Views` deal with everything graphical; they display aspects of their models. `Controllers` contain the interface between their associated models and views and the input devices. They send messages to the model and provide the interface between the model and its related views [7]. The MVC style provides flexibility that decreases code complexity and brings modularity to the system [1]. Thus, the platform obtains, on some level, benefits such as code reuse, high cohesion (due to MVC logical grouping), and joint development (because of the code modularity and independence).

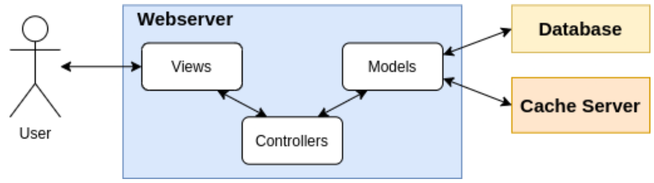

Figure 2: Basic software components of the platform

Figure 2 shows the software components of the developed platform. The `Webserver` is the component where we implemented the MVC, using the Ruby on Rails framework, with the domain logic that encodes the business rules. By itself, Rails does not provide storage for the model's data, requiring connection to an external relational `Database`. With a large dataset, queries to the database can be a performance-intensive task, causing responsiveness issues on the platform. One solution to this issue is using a `Cache Server`, where the answers for the most frequent queries can be stored to speed up data access. This solution avoids wasting time on work that has already been done, increasing the system's overall performance and responsiveness. Currently, our platform uses *PostgreSQL* as a database management system and *Redis* as a cache server.

The SIH-SUS dataset structure may subtly differ from region to region, so we also need a *database generalizer* to receive the different SIH-SUS datasets and make their visualization possible despite these differences. The importation of a SIH-SUS dataset characterizes this generalization into the platform. A dataset is imported into the platform through a CSV file containing some predefined columns (such as the hospitalization admission date, the health facility code, the patient age, and 26 other fields).

The platform does not store identifiable patient information such as name and exact residence location; it only stores his census tract. However, it can be possible to identify a patient with a rare disease through his census tract. For example, a census tract can be a single building, and, within this building, there may be a single case of a given disease. To overcome this exposure, we use the *Devise* gem from Ruby on Rails for user authentication to control data access in the platform. It uses the *Bcrypt* password hashing algorithm and saves encrypted passwords directly on the database. This way, the platform securely stores data avoiding unauthorized access.

In terms of our SIH-SUS dataset, we are dealing with more than half a million records, and such a large number can also become an obstacle to data analysis from both visualization and response time perspectives. Classifying or grouping the data into a set of categories is a way to solve the visualization problem. Our platform provides a spatial visualization through clusters since the data are displayed on a geographic map. With the computational clustering approach, the platform supports public health professionals in analyzing and understanding geographic patterns. The platform addresses the response time issue by employing denormalization techniques [9] and caching to *Redis* the generated *JavaScript* code for each data point.

[4] https://gitlab.com/interscity/health-dashboard

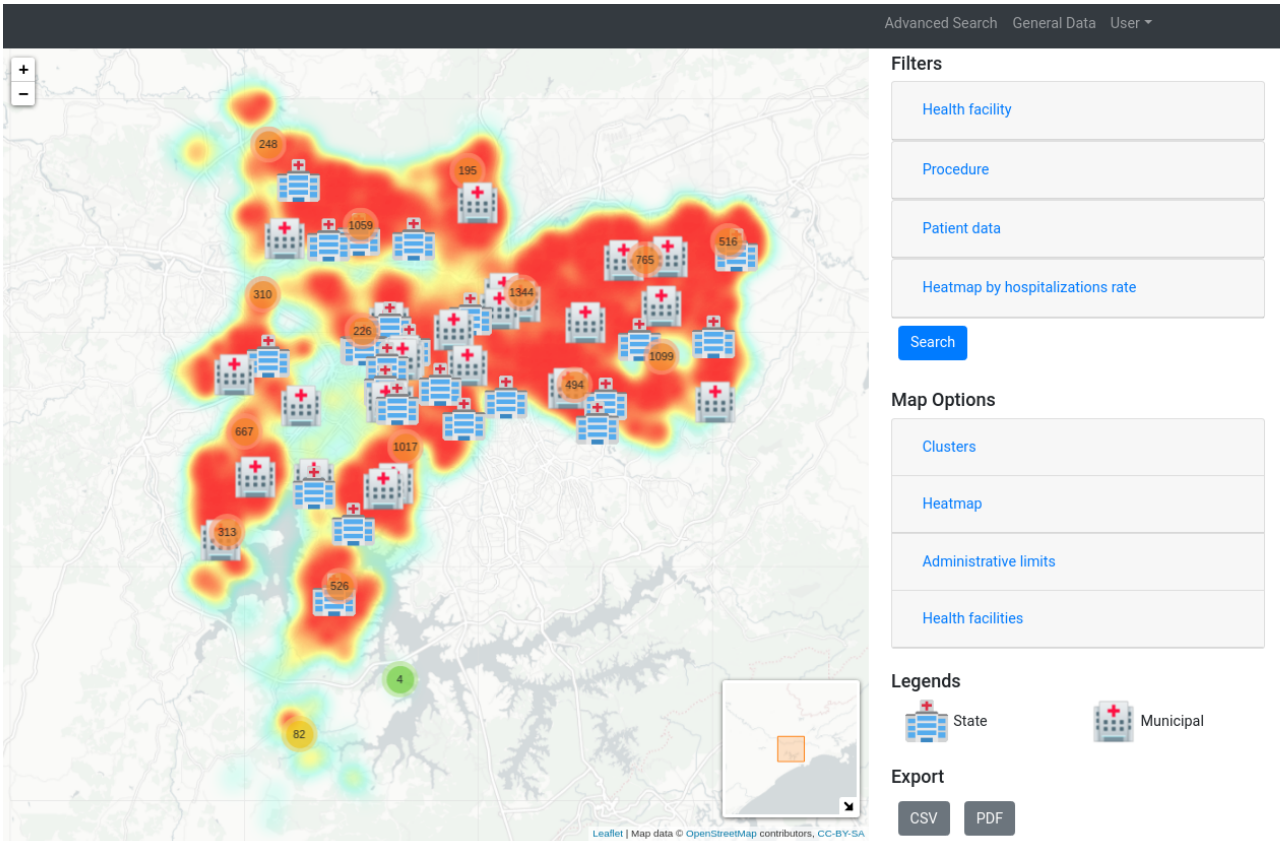

Figure 3: An Advanced Search displaying all hospitalizations due to a specific disease

6 Features

The user can view a map with information about hospitalizations and health facilities from the SIH-SUS dataset of our partner in the platform. It is possible to import a SIH-SUS dataset from any other region since it fits the database generalizer script requirements.

The platform is a web application divided into two different pages for different purposes. The main page is called `Advanced Search` and is presented in Figure 3. It contains the *Procedure* model's view, featuring the patients' geo-anonymized locations, the health facilities' locations, and the region's hospitalization concentration. Inside the `Advanced Search` page, we can find the right side menu with the filters – Health facility, Procedure, Patient Data, and Heatmap by hospitalizations rate – to control the map visualization and the export buttons.

The three first filter sections refer to the data itself, while the fourth selects the total population used to compute the heatmap weighted by each census tract population, enabling to restrict it by gender and race. The data referring to filter fields from the first three sections enables a database query interface. It allows the visualization of the hospitalization information on the map by choosing any database attribute as a filter. The first section, Health facility, contains fields such as the health facility name, administration, and beds. The second, the hospitalization data (also called Procedure) section, includes the patient diagnosis and hospitalization date. The third section, Patient data, includes the patient's age, gender, and educational level. We can search by these composite filters and receive the hospitalization data that satisfies the query. This engine complies with the need to access the information in a large dataset. The platform supports a space-time pattern identification through the selected filters, *i.e.*, it enables the analysis of geographic and temporal variations of the hospitalizations over the city.

The data are visualized on a map through a heatmap and clusters representations. The colorful heatmap represents the intensity of the hospitalizations in each region. The clusters represent the quantity of these hospitalizations. The data are grouped by distance-based clustering over the location of the health facility where the hospitalization occurred. An orange cluster contains more than 100 group records, a yellow cluster contains less than 100, and a green cluster contains all the hospitalization records of a given census tract. The map also supports visualization of administrative limits, *e.g.*, city limits and sub-prefectures, and health facilities percentiles circles.

Figure 4 displays another example from the `Advanced Search` page. It depicts hospitalizations from two different health facilities: a highly-specialized hospital from the University of São Paulo that provides high-complexity services and a local, small hospital. The concentric circles show the areas in which the patients of each of these hospitals reside. The blue, green, and red circles show the region in which 25, 50, and 75% of the patients reside. This functionality allows for the analysis of the geographical coverage of different health facilities. In this example, the university hospital serves the entire city, and the other one serves only close neighborhoods. Have in mind that we are dealing with patients coming majorly from middle- and low-income strata,

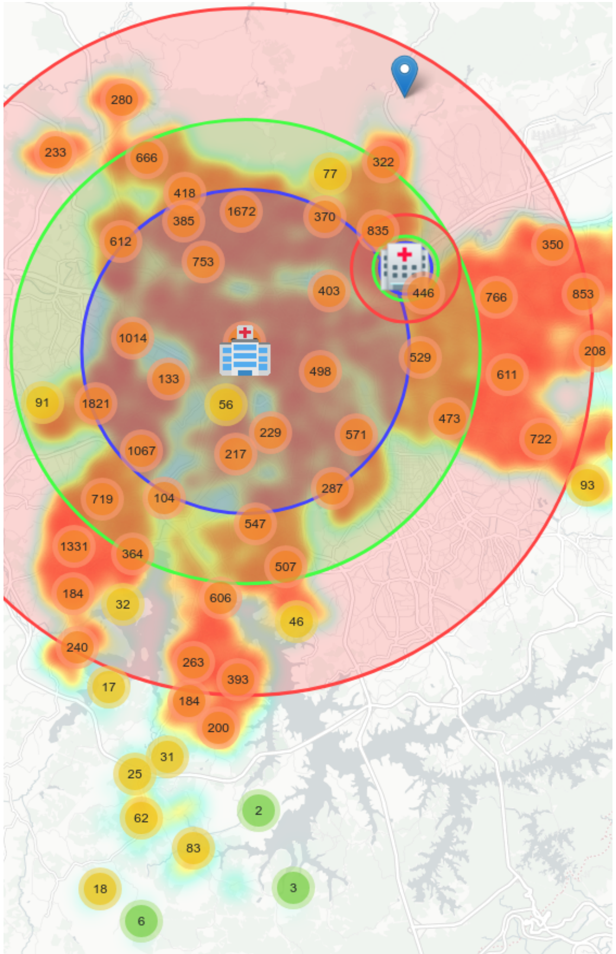

Figure 4: An Advanced Search displaying all hospitalizations from two specific health facilities with their respective percentile circles

and they live in a large metropolis where transportation can be a real problem. Thus, analyzing the location of the facilities and the residence of their patients is a very relevant issue.

The platform also contains a page called `General Data`, composed of a wide variety of charts. Each database attribute is displayed in a chart on this page, subject to the `Advanced Search` page's selected filters. We designed all these features to meet functional requirements identified in meetings with the healthcare experts from our partner, aiming to provide a proper public health management tool.

The `General Data` page's top shows the database info and all the current applied filters (Figure 5). The Ranking section contains tables ordered by the number of hospitalizations in different data attributes. This feature allows the observation of which health facilities have more hospitalizations, as illustrated in Figure 6.

The Health Facilities section presents charts informing the number of hospitalizations by health facility type and the number of beds in each health facility. The Hospitalization section (Figure 7) shows hospitalizations by the number of days in the hospital. It is also available in this section the proportion of hospitalization types and patient gender, age, and race information.

Figure 8 shows the Distances section, containing data related to the distances between patients' homes and health facilities where they are serviced [5]. The mean distance traveled for each specialty can be found in this section. There is also another distance-related section with a chart for each health facility presenting the proportion of hospitalizations in four distance groups for each one of the specialties. Considering that a low-income patient may spend up to 1 or 2 hours in a crowded bus to reach the proper health facility, this is a significant concern.

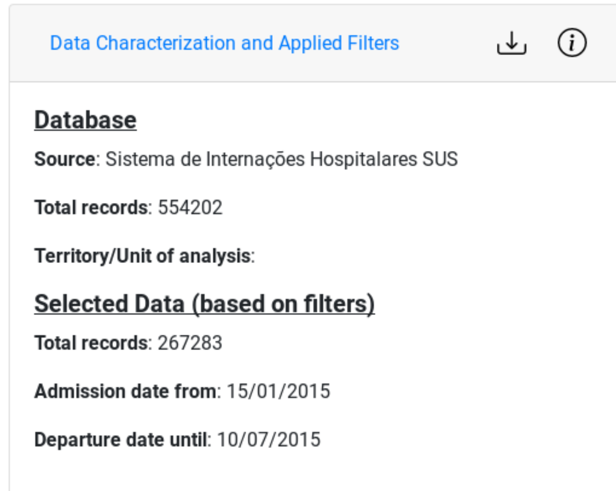

Figure 5: Database info and applied filters on `General Data` page

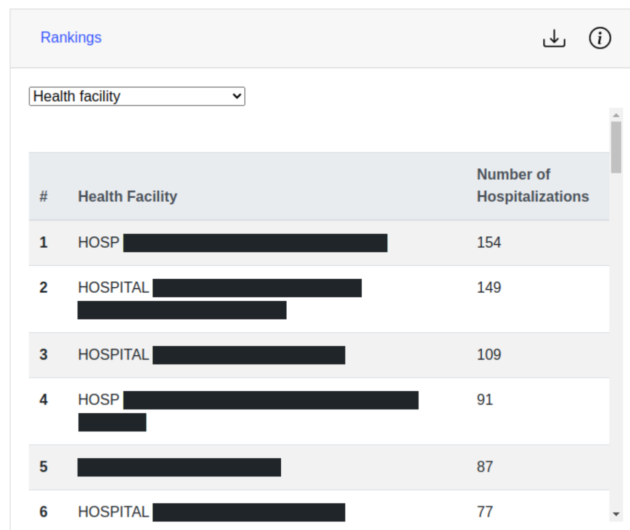

Figure 6: Ranking of health facilities (anonymized here) with more hospitalizations due to a specific disease

Hospitalization data in four different administrative levels are displayed on choropleth maps in the Territories section. Figure 9 shows one of these maps. The Dynamic Chart section produces individual charts based on a list of relevant categories given by our partner, including treemaps with the number of hospitalizations by diagnosis in the ICD-10.

[5]Using the Open Source Routing Machine (OSRM) software, we calculated the distance via city streets to better estimate how much a patient traveled to get medical assistance.

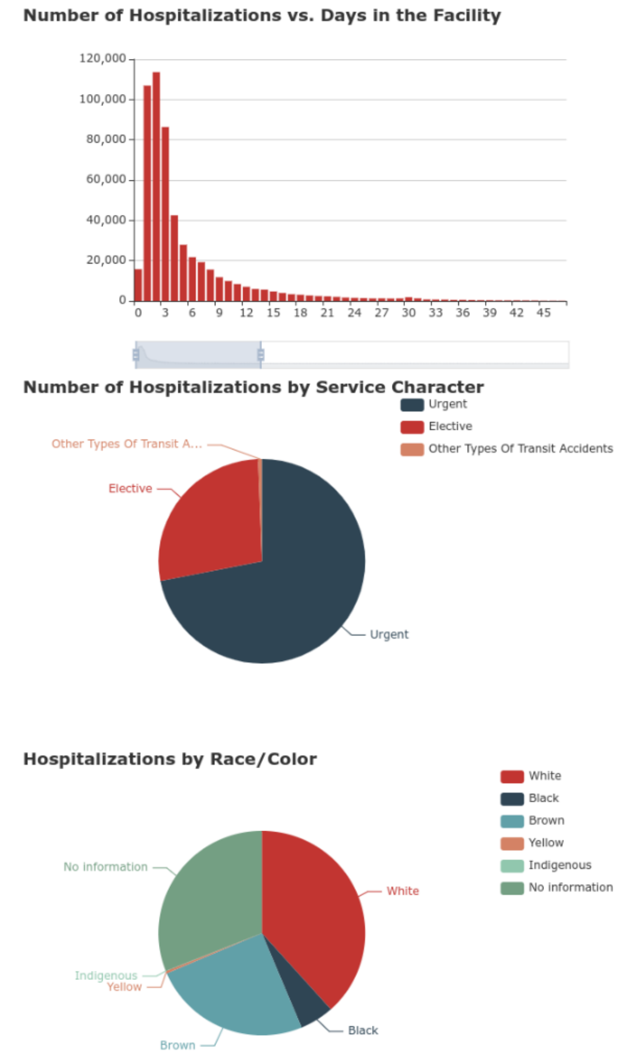

Figure 7: Hospitalization section with information of patients and hospitalizations due to a specific diagnosis

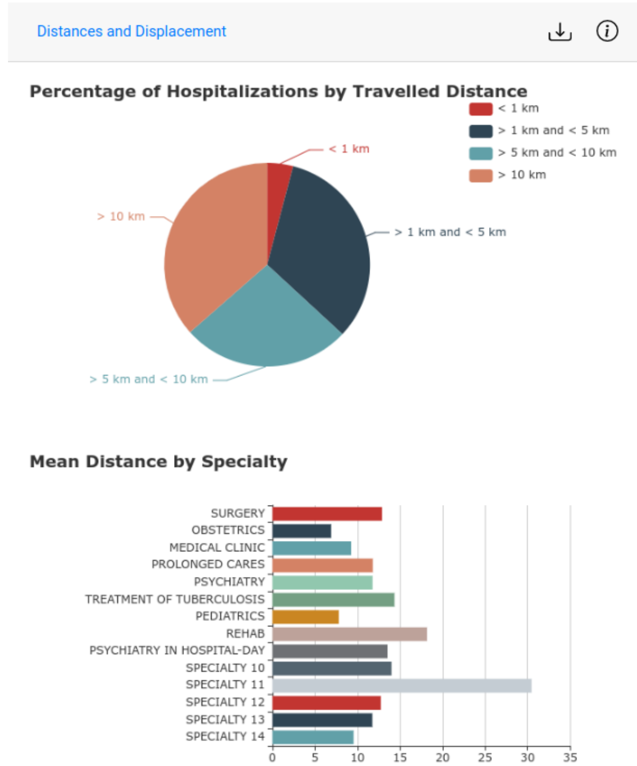

Figure 8: Partial view of the Distances section

As shown in Figure 10, the Descriptive Statistics section contains a table with the sum, the minimum, the maximum, the mean, and the standard deviation of some database attributes like patient age, days in the hospital, or ICU days. Finally, the Census Tract section shows a table with patient data distribution by gender and race for each census tract (the smallest area division in the platform).

Both pages support exports of the view. In the `Advanced Search` page, users can download the resulting data as a CSV file, a filtered subset of the original dataset, or a PDF file containing the map view, applied filters, legend, and source. In the `General Data` page, users can export each section as a PDF file, including all the selected section charts and their source.

Within our platform, healthcare expert users can evaluate the health situation in specific regions, the hospitalizations distribution on the map over time, area, diagnosis, and many other variables, according to the requirements and homologation tests from the Brazilian public health officials of our partner. The users can also be aware of the distance traveled by patients to be hospitalized. The use of the `Advanced Search` page to analyze the distribution of hospitalizations by a specific health facility is another example. An expert user can retrieve this information over different periods and analyze the evolution of a specific scenario.

7 Conclusion and Future Work

Our platform provides a large number of possibilities for combinations of information. The `Advanced Search` and `General Data` pages offer many variables to make those combinations, including patient diagnosis, age, race, gender, and region. As a result, our platform can produce pieces of evidence to establish effective healthcare policies.

From the technical point of view, our solution does not depend on sophisticated computing infrastructure and can be extended to other SUS data types. In the current version, we are evaluating its scalability from a dataset with 554,202 hospitalizations records. Our partner is part of the development process, and its healthcare officials test and approve all features as expert users in the homologation environment before we generate a new platform version.

Our next step is to include data from other years into the platform, provided by our Health Secretariat partner, and other datasets obtained directly from DATASUS. In this direction, our partner recently shared a new dataset covering 565,061 hospitalization records from 2018 and 577,905 from 2019, totaling 1,142,966 hospitalization records. This newer dataset will be incorporated soon into our platform.

After adding more data to the platform, we will then work on performance benchmarks and improvements based on these results. Finally, a careful analysis of the feedback provided by professionals from the Health Secretariat during their daily use of the platform for a few months will provide the basis for refinements and the development of an enhanced version of the system. This evolution is essential

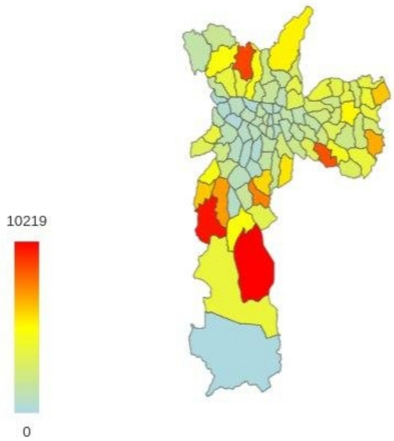

Figure 9: Number of hospitalizations per district

Descriptive Statistics

Variable	Count	Sum	Minimum	Maximum	Mean	Standard Deviation
Age	999	35991.0	0	96	36.02	22.07
Hospitalization days	999	5411.0	0	112	5.41	8.93
Service value	999	72878.48	0.0	51873.75	72.95	1649.97
ICU daily cost	999	442.0	0	89	0.44	3.6
IMCU daily cost	999	0.0	0	0	0.0	0.0
General daily cost	999	4969.0	0	112	4.97	7.88

Figure 10: Partial view of the Descriptive Statistics section

to enable the platform to handle the datasets from all Brazilian states, with public health data from the more than 150 million patients of the Brazilian National Health System. If we are successful in this, HealthDashboard may become a useful open-source tool contributing to a better healthcare in the nation.

Acknowledgements

This project is funded by CNPq (proc. 465446/2014-0), CAPES, and FAPESP (proc. 14/50937-1 and 15/24485-9).

References

[1] E. Gamma, R. Helm, R. Johnson, and J. Vlissides. *Design Patterns: Elements of Reusable Object-Oriented Software.* Addison-Wesley Longman Publishing, USA, 1995.

[2] H. A. Girade, O. M. d. Fonseca, J. P. Pereira, M. I. P. Costa, A. K. Lima Filho, L. L. Esmeraldo, and C. A. G. Tibiriçá. Conecta SUS Zilda Arns Neumann - Centro de Informações e Decisões Estratégicas em Saúde: contribuição para a ampliação da capacidade gestora em saúde do Estado de Goáis. In *XV Congresso Brasileiro de Informática em Saúde (CBIS)*, pp. 130–131, 2016.

[3] J. T. Hertz, T. Fu, J. R. Vissoci, T. A. H. Rocha, E. Carvalho, B. Flanagan, L. de Andrade, A. T. Limkakeng, and C. A. Staton. The distribution of cardiac diagnostic testing for acute coronary syndrome in the brazilian healthcare system: A national geospatial evaluation of health access. *PloS one*, 14(1):e0210502, 2019.

[4] A. Hunt and D. Thomas. *The Pragmatic Programmer: From Journeyman to Master.* Addison-Wesley Longman Publishing, USA, 2000.

[5] G. Iyer, S. DuttaDuwarah, and A. Sharma. Datascope: Interactive visual exploratory dashboards for large multidimensional data. In *2017 IEEE Workshop on Visual Analytics in Healthcare (VAHC)*, pp. 17–23, 2017.

[6] F. Kon, K. Braghetto, E. Z. Santana, R. Speicys, and J. G. Guerra. Toward smart and sustainable cities. *Commun. ACM*, 63(11):51–52, Oct. 2020.

[7] G. E. Krasner and S. T. Pope. A cookbook for using the model-view controller user interface paradigm in smalltalk-80. *J. Object Oriented Program.*, 1(3):26–49, Aug. 1988.

[8] J. Marinho de Souza, R. Giuffrida, A. P. M. Ramos, G. Morceli, C. H. Coelho, and M. V. Pimenta Rodrigues. Mother-to-child transmission and gestational syphilis: Spatial-temporal epidemiology and demographics in a brazilian region. *PLoS Neglected Tropical Diseases*, 13(2):e0007122, 2019.

[9] G. L. Sanders and S. Shin. Denormalization effects on performance of RDBMS. In *Proceedings of the 34th Annual Hawaii International Conference on System Sciences*, pp. 9–pp. IEEE, 2001.

[10] E. F. Z. Santana, A. P. Chaves, M. A. Gerosa, F. Kon, and D. S. Milojicic. Software platforms for smart cities: Concepts, requirements, challenges, and a unified reference architecture. *ACM Computing Surveys*, 50(6):78:1–78:37, 2017.

[11] W. Scacchi, J. Feller, B. Fitzgerald, S. Hissam, and K. Lakhani. Understanding free/open source software development processes. *Software Process: Improvement and Practice*, 11(2):95–105, 2006.

[12] A. Solanas, C. Patsakis, M. Conti, I. S. Vlachos, V. Ramos, F. Falcone, O. Postolache, P. A. Pérez-Martínez, R. Di Pietro, D. N. Perrea, et al. Smart health: a context-aware health paradigm within smart cities. *IEEE Communications Magazine*, 52(8):74–81, 2014.

[13] A. Sopan, A. S.-I. Noh, S. Karol, P. Rosenfeld, G. Lee, and B. Shneiderman. Community health map: A geospatial and multivariate data visualization tool for public health datasets. In *Government Information Quarterly*, vol. 29, pp. 223–234, 2012.

[14] K. J. Sullivan, W. G. Griswold, Y. Cai, and B. Hallen. The structure and value of modularity in software design. *SIGSOFT Softw. Eng. Notes*, 26(5):99–108, Sept. 2001.

[15] M. Wen, R. Siqueira, N. Lago, D. Camarinha, A. Terceiro, F. Kon, and P. Meirelles. Leading successful government-academia collaborations using floss and agile values. *Journal of Systems and Software*, 164:110548, 2020.

[16] C. C. Yang, G. Leroy, and S. Ananiadou. Smart health and wellbeing. *ACM Transactions on Management Information Systems (TMIS)*, 4(4):15, 2013.

[17] A. Yañez, J. Duggan, C. Hayes, M. Jilani, and M. Connolly. Pandemcap: Decision support tool for epidemic management. In *2017 IEEE Workshop on Visual Analytics in Healthcare (VAHC)*, pp. 24–30, 2017.

[18] Z. Yu, K. Pepe, G. Rust, J. E. Ramirez-Marquez, S. Zhang, and B. Bonnet. Patient-provider geographic map: An interactive visualization tool of patients' selection of health care providers. In *2017 IEEE Workshop on Visual Analytics in Healthcare (VAHC)*, pp. 1–8, 2017.

[19] W. Zhao, D. Borland, A. E. Chung, and D. Gotz. Visual cohort queries for high-dimensional data: A design study. In *2018 Workshop on Visual Analytics in Healthcare (VAHC)*, pp. 1–8, 2018.