Problem-driven visualization design of health and pollution big data

Design orientado a problemas para visualização de big data em saúde e poluição

Júlia Rabetti Giannella, Ximena Illarramendi, Renato Campos Mauro, Lucas Barcellos Oliveira, Igor Falconieri Santiago, Claudio Esperança, Doris Kosminsky

public health, big data, visual exploration, information visualization, decision making

We present the design of a big data visualization application aimed at investigating the potential relationships between air pollution and perinatal health. We integrated data from singleton pregnancies in Brazil over a seven-year period and records of fine particulate matter ($PM_{2.5}$) levels. Our methodology combines Design by Immersion and co-design, engaging specialists from various domains in problem-driven visualization design. The interactive visualization allows exploratory queries and comparative analyses based on medical birth certificate data restructured according to estimated conception dates. The collaboration of various domain experts was essential for leveraging complex data for an informed decision-making visualization.

saúde pública, big data, exploração visual, visualização de informação, tomada de decisão

Este artigo apresenta o design de uma ferramenta de visualização de big data voltada para investigação das possíveis relações entre poluição atmosférica e saúde perinatal. Foram empregados dados públicos brasileiros de gestações únicas e de material particulado fino ($MP_{2.5}$) no decorrer de um período de sete anos. Nossa metodologia combinou Design by Immersion e co-design com a colaboração de especialistas de vários domínios para o design de uma visualização direcionada por problema. A visualização interativa resultante permite a exploração e comparação de dados de nascimento segundo datas de concepção estimadas, sem sugerir causalidade. Ao longo do processo de design, a colaboração com especialistas de diferentes áreas do saber mostrou-se essencial para a compreensão dos dados e o design de uma visualização de big data para tomada de decisão informada.

1 Introduction

There is increasing evidence of air quality and climate change effects on health (Liu et al., 2022; Romanello et al., 2022). The current environmental crisis poses a substantial health risk to vulnerable populations, in particular, to pregnant women, the developing fetus, and newborns (International Federation of Gynecology and Obstetrics [FIGO], 2020). Understanding the short- and long-term impact of air pollution on pregnancy outcomes and on children is crucial for the health and well-being of all generations (Di Renzo et al., 2015). Increasing evidence suggests that environmental exposure worsened by climate change, including exposure to fine particulate matter less than 2.5 µm in diameter ($PM_{2.5}$), leads to poor pregnancy outcomes such as preterm birth, low birth weight, stillbirth, infertility, and miscarriage (Pedersen et al., 2013; Klepac et al., 2018).

Studies on pollution's effect on health frequently focus on particular populations or specific geographical and socioeconomic conditions. These studies assess individual exposure to pollution, which requires statistical analysis of pollution and subject distribution over time and space. However, our initiative focuses on understanding the effects of air pollution on singleton pregnancies registered in Brazil over a seven-year period from a different perspective.

Brazil, a continental country with a population of 203 million (Brazilian Institute of Geography and Statistics [IBGE], 2022), possesses government-maintained freely available databases on health and the environment. The health data is available from the Brazilian Ministry of Health (MoH) through DATASUS[1] in aggregated form and information cannot be traced back to individual cases. Climate and environmental data is also freely available from the Ministry of Science, Technology, Innovation and Communication. These datasets are not linked, thus, when using them for assessing air pollution exposure during pregnancy must be estimated based on air quality conditions in the municipality where the mother lived during pregnancy.

1 DATASUS – Informatics Department from the Brazilian National Health System (SUS – Sistema Único de Saúde).

This paper describes the development of a visualization application (Figure 1) for public health specialists and decision-makers to explore data sets on $PM_{2.5}$ and perinatal health. The goal was to enhance understanding of the factors influencing pollution's impact on pregnancies. Designing

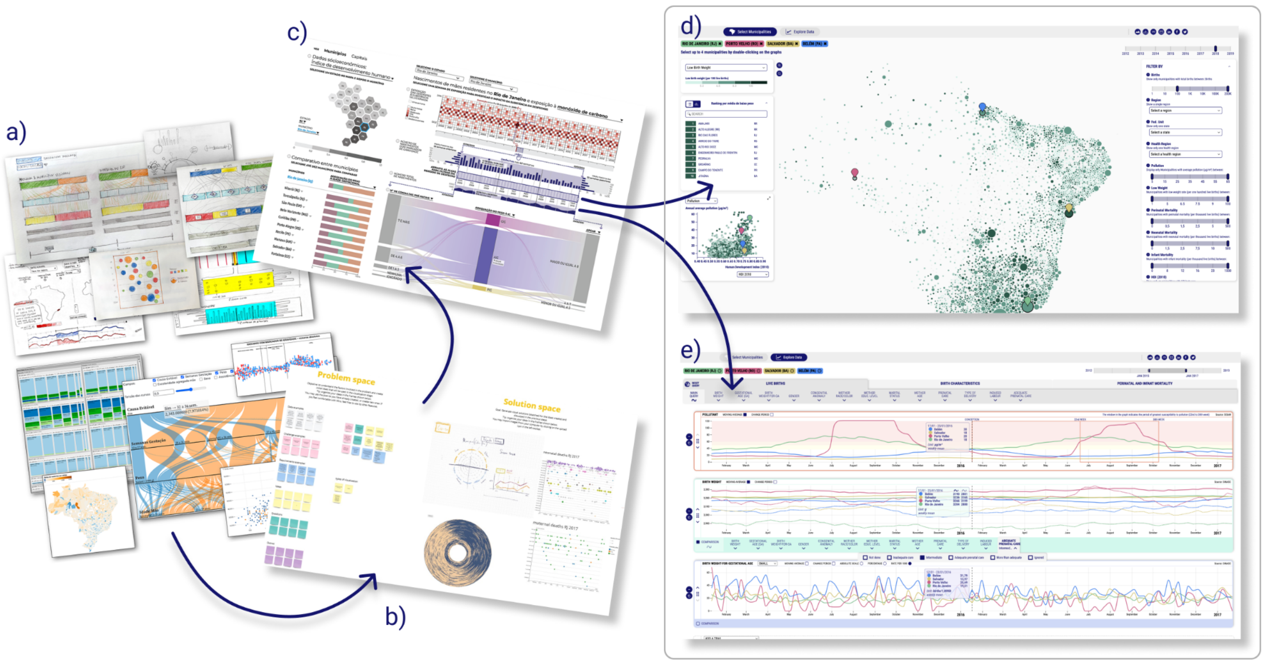

Figure 1 From concepts to the final visualization application.

visualizations based on such complex data is a 'wicked problem" (Rittel & Webber, 1973), due to the variety of contributing factors like socioeconomic conditions, diseases, and prenatal care. To address this challenge, we formed a transdisciplinary team and used the design by immersion approach (Hall et al., 2020) along with co-design participatory methods (Sanders & Stappers, 2008). The final product is intended to generate insights about the issue.

2 Related work

2.1 Information dashboards

The analysis of how environmental factors affect human health depends on the measurement precision of location and duration. Cartograms are frequently used to illustrate the distribution of one or few variables in space at a specific time (Gao et al., 2019). In contrast, time series visualizations typically use Cartesian graphs such as line charts or scatter plots, with time on the *X*-axis and other variables on the *Y*-axis (Fang et al., 2020). Other designs for time series include spiral plots (Carlis & Konsta, 1998), calendar-based views (Wijk & Selow, 1999), and ThemeRiver charts (Havre et al., 2000).

Combining maps with conventional time series plots using the "small multiples" approach (Pena-Araya et al., 2020) allows visualization of both spatial and temporal variables, but this approach is limited to a relatively small number of geographical points and/or summarized time-based charts. Another way to add a time dimension to visualizations is through animation and/or interaction, creating dynamic charts where temporal changes are shown through rapidly changing images (Robertson et al., 2008). Interaction in spatio-temporal exploration is more effective when used in a coordinated view context (Shimabukuro et al., 2004). Visualizations that combine multiple charts are often called "dashboards" (Sarikaya et al., 2019). While dashboards are typically designed to fit on a single screen for easy monitoring (Few, 2006), they can support tabbed layouts for switching between pages or frames, each providing different decision-making components or context.

2.2 Visualizing health and environment data

In this section, our focus is on publications detailing custom interactive tools specifically designed for visualizing the relationship between health and environmental data. We identified several articles meeting these criteria, with eight considered particularly relevant. We summarize the findings of these articles in Chart 1.

Regarding the health-related variables represented, preterm births and women's health were only found in Jankowska et al. (2019), which makes such tool the closest to the present work, although the authors used primary health-related variables obtained from linked hospital discharge, birth certificates, and infant death records.

Chart 1 Interactive visualizations tools in the Literature.

Source	Data structures	Environmental data variables	Location (scale)
Hanrahan et al. (2004)	choropleth maps	$PM_{2.5}$	Wisconsin (state)
Parmanto et al. (2008)	choropleth maps, bar charts	$PM_{2.5}$, water sources, natural disasters	Indonesia (country)
Rouillon et al. (2017)	geographic maps	Mn, Cu, Zn, As and Pb concentrations in soils from urban vegetable gardens	Sydney (countie)
Jankowska (2019)	choropleth maps, heatmaps	$PM_{2.5}$, water pollution, pesticides,	Fresno (countie)
Yang et al. (2019)	bar charts, line graphs	$PM_{2.5}$	Taiwan (country)
Alonso et al. (2022)	choropleth maps	$PM_{2.5}$, temperature, humidity	Balearic Island (region)
Chaudhuri et al. (2022)	choropleth maps, dot maps, flow maps, line graphs and scatterplots	$PM_{2.5}$, temperature, humidity	Catalonia (region)
Khreis et al. (2023)	geographic maps, pie and bar charts	Traffic-related air pollution effects and pollutants	World

2.3 Problem-driven visualization design

Various methodologies exist for problem-driven visualization design, each offering unique guidelines and approaches. For instance, Sedlmair et al. (2012) suggest a methodology that involves the careful selection and assignment of roles to experts before initiating design studies. McCurdy et al. (2016) emphasize mutual learning among team members and recommend assigning collaborative roles through deliberate planning.

In contrast, Design by Immersion (Hall et al., 2020) proposes a methodology where researchers immerse themselves in another domain work, fostering transdisciplinary experiences and interactions to generate new knowledge. Additionally, the co-design approach (Sanders & Stappers, 2008) involves non-designers working together in the design development process. These methods have gained traction across various domains, particularly in health research where all participants contribute to the co-creative process (Slattery et al., 2020).

3 Methods

We adopted the Design by Immersion methodology (Hall et al., 2020), which typically involves case studies that integrate visualization with other domain disciplines, like literature or chemistry. However, our approach extended beyond dual immersion, incorporating three primary disciplines: visualization, public health, and environmental science.

A core team, referred to as the infovis team, consisted of five specialists in visualization (from computer science and visual design) and two public health experts, one of whom had knowledge on environmental issues. The infovis team met weekly online for 24 months and discussed every step of the process, from sketches to final design proposals.

The domain collaborators, 49 experts from six different knowledge domains, were involved depending on the phase of the design process (Table 1).

Table 1 Domain collaborators participating in Phases 3 and 4: Online meetings and workshops (ws).

Activity	Data Engineers	Communic. or Visual Designers	Health Managers	Stats.	Environ. scientist	MD or public health
1st Domain Expert ws	4	13	1	1	0	0
2nd Domain Expert ws	4	10*	0	2	3	4
Final meetings	0	1	2	0	0	6
Expert evaluation	0	4*	5	0	0	1

(*) Designers acted as workshop facilitators. Communic. = communication, Stats. = Statisticians, MD = Medical Doctor.

The collaboration activities followed the notion of co-design (Slattery et al., 2020) and occurred through seven iteration cycles of workshops or meetings. Except for the expert evaluation (which will not be covered in this paper), all design workshops and activities were conducted remotely, due to the isolation policies required by the COVID-19 pandemic.

4 Design process

We established four phases for the design elaboration. The duration of each phase was 7, 3, 12, and 2 months, respectively.

4.1 Phase 1. Data understanding

The original datasets used in the project are summarized in Table 2. The DATASUS datasets are complex and have high-dimensional structure. We used versions extracted, transformed, and loaded by the PCDaS[2] of the Oswaldo Cruz Foundation (Fiocruz), MoH.

The air pollution data was available from the Health Integrated Environmental Information System (SISAM), National Institute for Spatial Research (INPE) for the period between 2000 and 2019. The dataset was developed for studying the environmental effects on human health and comprises daily measurements of meteorological data and air pollutant levels, for 5570 Brazilian municipalities (Setzer & Ferreira, 2021). Lastly, we used the Human Development Index (HDI) computed for 2010 by the Institute for Applied Research on Economics (IPEA) and covering 5565 municipalities.[3]

The infovis team utilized concept maps (McKenna et al., 2014) to understand the databases, synthesize data relationships, and establish necessary indices and parameters. Concept sketching (Figure 1a) was employed to test ideas. The most promising sketches were developed as

2 Data Science Platform Applied to Health.

3 The five municipalities that were created in 2013 – Pescaria Brava (SC), Balneário Rincão (SC), Mojuí dos Campos (PA), Pinto Bandeira (RS) and Paraíso das Águas (MS) – are not included in our visualizations.

Table 2 Databases, sources, and basic characteristics.

Source	Database*	Number of records	Attributes
PCDaS	SINASC – Sistema de Informação sobre Nascidos Vivos	74,398,079	141
PCDaS	SIM – Sistema de informação sobre Mortalidade	27,902,867	178
PCDaS	DOFET – Declaração de óbito fetal	857,664	168
INPE	SISAM – Sistema de Informações Ambientais Integrado à Saúde	162,720,232	19
IPEA	HDI	5,565	2

(*) SINASC – Medical Birth Certificates; SIM – Medical Death Certificates for Newborns and Mothers; DOFET – Fetal death certificates.

JavaScript prototypes (Figure 1c) for testing with real data. This approach, both generative and evaluative, allowed the infovis team to address the Domain Problem (DP) while creating visual solutions.

4.2 Phase 2 Addressing the design problem and challenge

After understanding the data, the infovis team could establish the following:

- **DP**: to visually explore the relationship between environmental pollution and maternal-newborn health;
- **Design Challenge (DC)**: to investigate and visually present the temporal relationship between pregnant women' exposure to pollutants and neonatal outcomes, without inducing false causal attributions between heterogeneous data from the unlinked databases,

Data Decisions (DD) were necessary to address the DC. The data was further aggregated and reduced in order to comply with the visualizations' requirements. The following five DD led to the initial visualization solution.

- **DD1 (Data aggregation)**: all data were aggregated by epidemiological week due to different temporal granularities of datasets;
- **DD2 (Single births selection)**: multiple pregnancies were excluded to reduce variability;
- **DD3 (Pollution and human health rates)**: the project adopted CONAMA's (National Council for the Environment) benchmark values for daily measurements $PM_{2.5}$ acceptable for human health;
- **DD4 (Conception Date)**: an estimated conception date was adopted by subtracting the gestation length in weeks from the date of birth to synchronize the observation of gestational periods and pollution;
- **DD5 (Birth weight variable)**: records without birth weight were excluded, determining the use of records from 2012 onward;

Causality is often perceptually linked to temporality (Dimara et al., 2020; Zuk & Carpendale, 2014), therefore DC influenced the visual structures, resulting in a dashboard with a calendar-based visualization for pollution data and a parallel sets graph (Bendix et al., 2005) for the health data. It also integrated a map and ranking view (Figure 1c).

4.3 Phase 3. Establishing the functional requirements

4.3.1 First domain expert workshop

The third phase began with a virtual co-design workshop to validate the first prototype and generate parallel prototypes (McKenna et al., 2014) for the DC. Participants had no access to previous design solutions. The workshop included an initial plenary session; an explanatory activity addressing the DP and DC; and a brainstorm activity in breakup rooms. Each group discussed and ideated the DP and DC using a 'talking aloud' protocol (Charters, 2003), followed by sketch generation on drawing board template to propose quick visual solutions (Figure 1b). The main outcome of the first workshop was the similarity between collaborators' design suggestions and the infovis team design, validating the chosen visual structures, like the calendar-based view that was considered suitable for this type of data because it communicates spatial distribution and density. It also reveals linear (over the years) and cyclical (over the months and seasons) temporal patterns.

Following the workshop results, the infovis team completed the implementation of the first visualization dashboard (Figure 1c), which was then explored in a second workshop.

4.3.2 Second domain expert workshop

The second workshop aimed to: a) assess effectiveness of the first prototype; b) gather insights for potential improvements; and c) evaluate the tool's utility in generating hypotheses regarding the relationship between perinatal health and pollution data. Following the Visual Data Analysis and Reasoning scenario (Lam et al., 2012), the evaluation focused on studying how the visualization tool supported the generation of actionable knowledge in this domain.

After a plenary session to provide an overview of the DP and DC, we conducted a 'think aloud' activity (Fonteyn et al., 1993), where two facilitators interacted with each specialist – one observed while the other took field notes. These sessions were recorded for further analysis of participant responses and narrative descriptions.

The analysis revealed several challenges faced by experts in visually exploring the data. One environmental expert expressed difficulty in relating pollution data to health outcomes, suggesting a desire for visual comparisons between cities with varying pollution levels. Another expert highlighted the clarity of the data but noted challenges in deriving meaningful messages from it, expressing a need for simpler, more intuitive visualizations akin to an '*x* and *y* graph'. The use of different visual structures to represent two time-related variables was deemed ineffective, indicating a need for greater flexibility in the dashboard's design. Experts struggled to identify patterns or visually correlate newborn health and air pollution data, leading to the abandonment of the calendar-based view and a parallel set chart concept. This workshop informed the formulation of five functional requirements for a revised design:

- **R1 (Comparison)**: allow to choose and compare municipalities of interest based on flexible criteria including environmental, socioeconomic, health, and geographic scenarios;
- **R2 (Temporality and filtering)**: enable visual exploration of the variables of interest over the data period, with functionalities for flexible filtering and aggregation at interactive rates;
- **R3 (Synchronization)**: present a large amount of secondary data of different natures in a synchronized way, thus revealing temporal proximity between different variables;
- **R4 (Stratification)**: compare population extracts with different health and socioeconomic profiles;
- **R5 (Usefulness)**: be useful and accessible to domain experts and public managers working with public health, while providing fast and fluid navigation.

We developed a high-fidelity wireframe and further prototyped the Exploration Trails view, which is a coordinated visualization with vertically-arranged cards containing line charts and filtering options. This design will be described in Section 5.

4.3.3 Domain collaborators meetings

During the third phase of the design process, several incremental design modifications were implemented and informally evaluated in three online meetings with members of the domain collaborators group. We conducted participatory observation and interviews during the meetings, with one member of the infovis team responsible for the field notes. These sessions were recorded and automatically transcribed.

The main DD resulting from these meetings was to resume the map and ranking views abandoned after the first design iteration (Fig .1c) to facilitate the selection of municipalities for later comparison (R1), which culminated in the creation of the select frame (map overview). Other minor decisions were: (a) include an 'adequate prenatal care' filter (R2); (b) design a *susceptibility window* feature in the pollution trail to highlight the period in which a fetus is most vulnerable to effects of $PM_{2.5}$ (R3); (c) include a filter panel in the 'select' frame (R2); (d) improve the interface of the 'select' frame to coordinate with the 'explore' frame more consistently and fluidly (R5); and (f) improve tool performance to mitigate delay issues (R5).

Beyond the comments related to the dashboard, one of the public health specialists validated the concentration levels of $PM_{2.5}$ deemed acceptable for human health and adopted in the project (DD3), and highlighted the huge incidence of air pollution in the Amazonian states as a consequence of the burning of forests to make room for agriculture and livestock. The intensity of this burning can produce effects in faraway municipalities insofar as the wind moves the smoke. This comment helped the infovis team understand the high incidence of $PM_{2.5}$ seen in small municipalities.

5 Final design

Phase 4 consisted in the final design implementation of the visualization Amplia Saúde (available for exploration at <ampliasaude.org>), a multi-page dashboard comprising two frames: The 'select' frame for municipality browsing and selection, and the 'explore' frame for detailed analysis. The frames are toggled using a button, allowing only one frame to be visible at a time. Additionally, the implementation includes: features for data download and social media sharing; along with a tutorial mode not discussed in this paper. Below we provide details of the select-and-explore design:

5.1 Select frame (Map Overview)

This frame provides an overview of Brazilian municipalities based on several variables (R1), including annual averages for $PM_{2.5}$; low birth weight rates; and perinatal, neonatal and infant mortality rates, as well as the HDI and the population data for 2018.[4]

The central visualization component is the Dorling cartogram (Nusrat & Kobourov, 2016) (Figure 2c), where municipalities are depicted as circles. The circle size is proportional to the municipality's population, and the color represents the scale of the selected variable (Figure 2e). This cartogram aids in identifying major cities and visually highlights sparsely populated regions (R4, R5) and areas with higher pollution levels.

Double clicking on circles selects municipalities for further analysis in the scatter plot and small-multiples bar charts (elements d, f, and g in Figure 2). These charts enable comparison of municipalities according to variables of interest over the 7-year data period (R2). A dropdown menu provides the selection of the variable of interest, which is measured through the color scale below (Figure 2e). A toggle button (Figure 2f) switches the small-multiples bar charts to a ranking list of municipalities based on the selected variable (see Figure 1). The filter panel (Figure 2h) allows users to display municipalities based on specific criteria (R2), and to restrict the view to a particular Region, State, or health region.[5]

4 A single year was used because the population has been mostly stable in the seven-year period covered by the tool.

5 The Brazilian MoH defines these as regions formed by border municipalities that share cultural, economic, and social identities, as well as communication networks and transportation infrastructure.

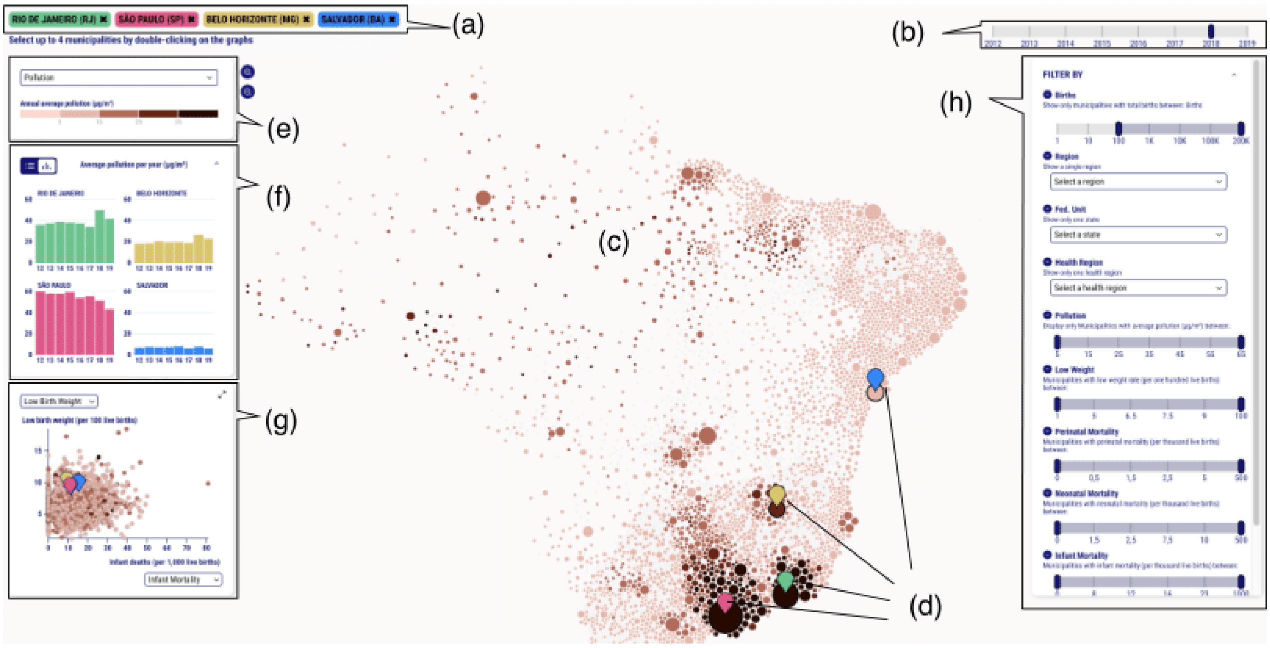

Figure 2 The Select Frame features.

5.2 Explore frame (Exploration Trails)

The Exploration Trails frame (Figure 3) comprises vertically-arranged cards (*trails*), each containing a line chart and auxiliary buttons and menus for filtering, grouping, and display options. Each trail displays weekly aggregated data (DD1) of one variable (*y*-axis) over time (*x*-axis) (Figure 3a). Hovering over a line activates a tooltip for on-demand details (Shneiderman, 1996) (Figure 3b). Variables can be selected from predefined sets based on exploration profiles, including live births, birth characteristics, and perinatal and infant mortality. $PM_{2.5}$ levels are shown by default as the first trail, with other variables reflecting the health exploration theme (Figure 3c).

Users can select up to four municipalities at a time to avoid cluttering the interface. Municipalities can be selected in this frame through the municipality menu (Figure 3a) or carried over from the 'select' frame. Each trail contains a moving average selector to smooth the line chart for better perception of seasonal behaviors. By default, the value for a given week is calculated as an average for that week and the six previous and subsequent weeks (Figure 3e).

Trails are mostly independent but have their *X*-axis aligned by default to ensure they cover the same time range, thus satisfying R3. This synchronization can be overridden on a trail-by-trail basis by checking the 'modify period' box, allowing comparisons between outcomes in different time periods (Figure 3f). Health-related trails include a comparison feature in the footer, complying with R1 (Figure 3g). Clicking this opens a local filtering menu to set health variables and socioeconomic filters for that specific trail, providing distinct perspectives for the analyzed data.

Pollution-related trails display benchmark ranges for acceptable, high, very high, and extreme values of $PM_{2.5}$ concentration as colored strips vertically stacked in the background (Figure 3h). Additionally, a *susceptibility window* is plotted over the graph, displaying the time interval between 22 and 38 weeks of pregnancy, considered to be the period in which the fetus is most susceptible to external effects caused by pollution (Figure 3i).

Trails can be reordered using drag handles (Figure 3j). New trails can be added by clicking the duplicate button of an existing trail (Figure 3k) or by adding a new trail button, thus enhancing insights from comparing different exploration trails.

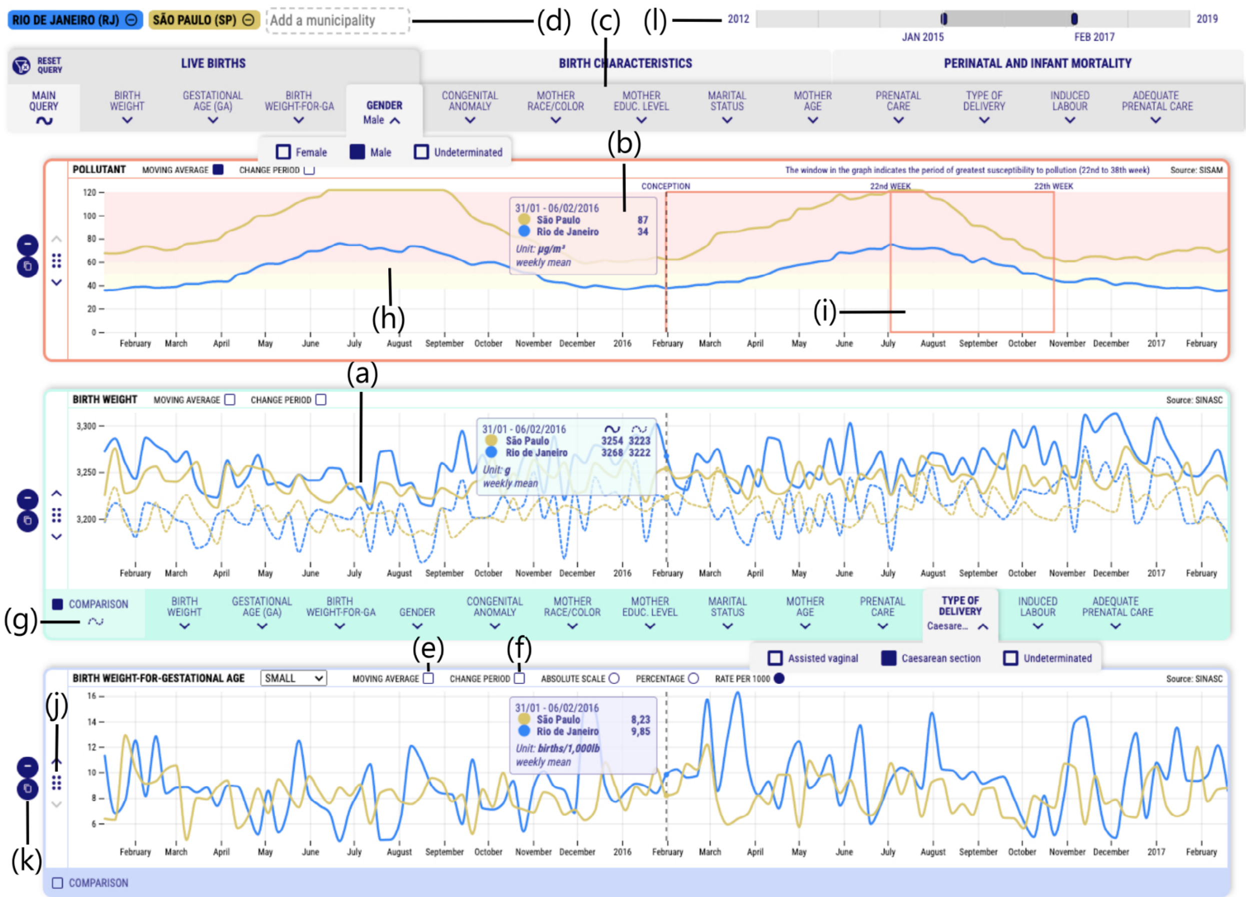

Figure 3 The Exploration trails visualization.

6 Conclusion and future work

The project described in this paper is deemed unique as it ventures into visualizing datasets not originally intended for combined analysis. The challenges in establishing concrete evidence linking air pollution to maternal and neonatal health issues have led to a complex landscape in data visualization design. Given the inherent uncertainty in this relationship, visualizations often aim to incorporate a vast array of data points to capture any potential influences on the process. Our approach involves integrating heterogeneous datasets and exploring myriad relationships to uncover potential correlations between air pollution and maternal and neonatal health outcomes. Thus, the visualizations strive to present a comprehensive view of the numerous factors at play, enabling researchers and stakeholders to gain insights despite the inherent uncertainties in the data. Our approach focused on presenting data views synchronized in both space and time. We identified the need for two levels of aggregation and synchronization: one for the spatial aspects and another for temporal aspects, achieved through the 'select' and 'explore' dashboard frames, respectively. The final tool design emerged after extensive collaboration with specialists from various domains. This collaborative effort was necessary due to the multifaceted nature of the relationships between health and environment, requiring a comprehensive approach. The final visualizations present heterogenous unlinked data in a manner that allows users to uncover insights relevant to their inquiries, favoring decision making. Along the process, feedback from collaborators supports this approach, with many noting that the dashboard has enriched their own exploration of the data.

Several limitations persist, largely stemming from the quality and coverage of the available data. The cessation of data processing by the SISAM team in 2019 is particularly unfortunate. Further improvements to the tool design are possible, such as enhancing the 'susceptibility window' in the pollution trail with additional derived information, like time convolutions with kernels covering selectable time periods.

Acknowledgment

This work was funded by the Bill & Melinda Gates Foundation (INV-027961) and Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq (445729/2020-1) and partially supported by Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro – FAPERJ (E-26/200.948/2022).

References

Alonso, E., Garcia-Navarro T., & Larburu, N. (2022). Visualization system for the impact of context open data in clinical settings. In J. Mantas, P. Gallos, E. Zoulias, A. Hasman, M. S. Househ, M. Diomidous, J. Liaskos, & M. Charalampidou (Eds.), *Advances in informatics, management and technology in healthcare* (pp. 128–131). Amsterdam: IOS Press. https://doi.org/10.3233/SHTI220678

Bendix, F., Kosara, R., & Hauser, H. (2005). Parallel sets: Visual analysis of categorical data. In J. Stasko & M. Ward (Eds.), *Proceedings of the IEEE Symposium on Information Visualization (InfoVis 2005)*, Minneapolis, MN (pp. 133–140). Los Alamitos, CA: Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/INFVIS.2005.1532139

Carlis, J., & Konstan J. (1998). Interactive visualization of serial periodic data. In E. Mynatt & R. J. K. Jacob (Eds.), *Proceedings of the 1998 11th Annual ACM*

Symposium on User Interface Software and Technology (UIST98), San Francisco, CA (pp. 29–38). https://doi.org/10.1145/288392.288399

Charters, E. (2003). The use of think-aloud methods in qualitative research: An introduction to think-aloud methods. *Brock Education Journal*, *12*(2). https://doi.org/10.26522/brocked.v12i2.38

Chaudhuri, S., Giménez-Adsuar, G., Saez, M., & Barceló, M. A. (2022). PandemonCAT: Monitoring the COVID-19 Pandemic in Catalonia, Spain. *International Journal of Environmental Research and Public Health*, *19*(8), 4783. https://doi.org/10.3390/ijerph19084783

Di Renzo, G. C., Conry, J. A., Blake, J., DeFrancesco, M. S., DeNicola, N., Martin Jr., J. N., McCue, K. A., Richmond, D., Shah, A., Sutton, P., Woodruff, T. J., van der Poel, S. Z., & Giudice, L. C. (2015). International Federation of Gynecology and Obstetrics opinion on reproductive health impacts of exposure to toxic environmental chemicals. *International Journal of Gynecology & Obstetrics*, *131*(3), 219–225. https://doi.org/10.1016/j.ijgo.2015.09.002

Dimara, E., Franconeri, S., Plaisant, C., Bezerianos, A., & Dragicevic, P. (2020). A task-based taxonomy of cognitive biases for information visualization. *IEEE Transactions on Visualization and Computer Graphics*, *26*(2), 1413–1432. https://doi.org/10.1109/TVCG.2018.2872577

Fang, Y., Xu, H., & Jiang, J. (2020). A survey of time series data visualization research. *IOP Conference Series: Materials Science and Engineering*, *782*(2). https://doi.org/10.1088/1757-899X/782/2/022013

Few, S. (2006). *Information dashboard design: The effective visual communication of data*. Beijing; Cambridge, MA: O'Reilly.

Fonteyn, M. E., Kuipers, B., & Grobe, S. J. (1993). A description of think aloud method and protocol analysis. *Qualitative Health Research*, *3*(4), 430–441. https://doi.org/10.1177/104973239300300403

Gao, P., Li, Z., & Qin Z. (2019). Usability of value-by-alpha maps compared to area cartograms and proportional symbol maps. *Journal of Spatial Science*, *64*(2), 239–255. https://doi.org/10.1080/14498596.2018.1440649

Hall, K. W., Bradley, A. J., Hinrichs, U., Huron, S., Wood, J., Collins, C., & Carpendale, S. (2020). Design by immersion: A transdisciplinary approach to problem-driven visualizations. *IEEE Transactions on Visualization and Computer Graphics*, *26*(1), 109–118. https://doi.org/10.1109/TVCG.2019.2934790

Hanrahan, L. P., Anderson, H. A., Busby, B., Bekkedal, M., Sieger, T., Stephenson, L., Knobeloch, L., Werner, M., Imm, P., & Olson, J. (2004). Wisconsin's environmental public health tracking network: Information systems design for childhood cancer surveillance. *Environmental Health Perspectives*, *112*(14), 1434–1439. https://doi.org/10.1289/ehp.7150

Havre, S., Hetzler, B., & Nowell L. (2000). ThemeRiver: Visualizing theme changes over time. In *Proceedings of the IEEE Symposium on Information Visualization (InfoVis 2000)*, Salt Lake City, UT (pp. 115–123). Los Alamitos, CA: Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/INFVIS.2000.885098

Instituto Brasileiro de Geografia e Estatística [Brazilian Institute of Geography and Statistics]. (2022). *Censo 2022*. https://www.ibge.gov.br/

International Federation of Gynecology and Obstetrics. (2020). *Climate crisis and health*. https://www.figo.org/climate-crisis-and-health

Jankowska, M. M., Yang, J.-A., Block, J., Baer, R. J., Jelliffe-Pawlowski, L. L., Flores, S., Pacheco-Warner, T., Costantino, A., Fuchs, J., Chambers, C. D., Newel, G. (2019). An online geographic data visualization tool to relate preterm births to environmental factors. *Preventing Chronic Disease*, 16. https://doi.org/10.5888/pcd16.180498

Khreis, H., Sanchez, K. A., Foster, M., Burns, J., Nieuwenhuijsen, M. J., Jaikumar, R., Ramani, T., & Zietsman, J. (2023). Urban policy interventions to reduce traffic-related emissions and air pollution: A systematic evidence map. *Environment International*, *172*. https://doi.org/10.1016/j.envint.2023.107805

Klepac, P., Locatelli, I., Korošec, S., Künzli, N., & Kukec, A. (2018). Ambient air pollution and pregnancy outcomes: A comprehensive review and identification of environmental public health challenges. *Environmental Research*, *167*, 144–159. https://doi.org/10.1016/j.envres.2018.07.008

Lam, H., Bertini, E., Isenberg, P., Plaisant, C., Carpendale, S. (2012). Empirical studies in information visualization: Seven scenarios. *IEEE Transactions on Visualization and Computer Graphics*, *18*(9), 1520–1536. https://doi.org/10.1109/TVCG.2011.279

Liu, D., Cheng, K., Huang, K., Ding, H., Xu, T., Chen, Z., & Sun, Y. (2022). Visualization and analysis of air pollution and human health based on cluster analysis: A bibliometric review from 2001 to 2021. *International Journal of Environmental Research and Public Health*, *19*(19). https://doi.org/10.3390/ijerph191912723

McCurdy, N., Dykes, J., & Meyer, M. (2016). Action design research and visualization design. In M. Sedlmair, P. Isenberg, T. Isenberg, N. Mahyar, H. Lam (Eds.), *Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization (BELIV '16)*, Baltimore, MD (pp. 10–18). New York: Association for Computing Machinery. https://doi.org/10.1145/2993901.2993916

McKenna, S., Mazur, D., Agutter, J., Meyer, M. (2014). Design activity framework for visualization design. *IEEE Transactions on Visualization and Computer Graphics*, *20*(12), 2191–2200. https://doi.org/10.1109/TVCG.2014.2346331

Nusrat, S., & Kobourov, S. (2016). The state of the art in cartograms. *Computer Graphics Forum*, *35*(3), 619–642.

Parmanto, B., Paramita, M. V., Sugiantara, W., Pramana, G., Scotch, M., & Burke, D. S. (2008). Spatial and multidimensional visualization of Indonesia's village health statistics. *International Journal of Health Geographics*, *7*(30). https://doi.org/10.1186/1476-072X-7-30

Pedersen, M., Giorgis-Allemand, L., Bernard, C., Aguilera, I., Andersen, A.-M. N., Ballester, F., Beelen, R. M. J., Chatzi, L., Cirach, M., Danileviciute, A., Dedele, A., van Eijsden, M., Estarlich, M., Fernández-Somoano, A., Fernández, M. F., Forastiere, F., Gehring, U., Grazuleviciene, R., Gruzieva, O., … Slama, R. (2013). Ambient air pollution and low birthweight: A european cohort study (ESCAPE). *Lancet Respiratory Medicine*, *1*(9), 663–754.

Pena-Araya, V., Pietriga, E., & Bezerianos, A. (2020). A comparison of visualizations for identifying correlation over space and time. *IEEE Transactions on Visualization & Computer Graphics*, *26*(1), 375–385. https://doi.org/ 10.1109/TVCG.2019.2934807

Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, *4*, 155–169. https://doi.org/10.1007/BF01405730

Robertson, G., Fernandez, R., Fisher, D., Lee, B., & Stasko, J. (2008). Effectiveness of animation in trend visualization. *IEEE Transactions on Visualization and Computer Graphics*, *14*(6), 1325–1332. https://doi.org/10.1109/TVCG.2008.125

Romanello, M., Di Napoli, C., Drummond, P., Green, C., Kennard, H., Lampard, P., Scamman, D., Arnell, N., Ayeb-Karlsson, S., Ford, L. B., Belesova, K., Bowen, K., Cai, W., Callaghan, M., Campbell-Lendrum, D., Chambers, J., van Daalen, K. R., Dalin, C., Dasandi, N., ... Costello, A. (2022). The 2022 report of the *Lancet* Countdown on health and climate change: Health at the mercy of fossil fuels. *The Lancet*, *400*(10363), 1619–1654. https://doi.org/10.1016/S0140-6736(22)01540-9

Rouillon, M., Harvey, P. J., Kristensen, L. J., George, S. G., & Taylor, M. P. (2017). VegeSafe: A community science program measuring soil-metal contamination, evaluating risk and providing advice for safe gardening. *Environmental Pollution*, *222*, 557–566. https://doi.org/10.1016/j.envpol.2016.11.024

Sanders, E. B.-N., & Stappers P. J. (2008). Co-creation and the new landscapes of design. *CoDesign*, *4*(1), 5–18. https://doi.org/10.1080/15710880701875068

Sarikaya, A., Correll, M., Bartram, L., Tory, M., & Fisher, D. (2019). What do we talk about when we talk about dashboards? *IEEE Transactions on Visualization and Computer Graphics*, *25*(1), 682–692. https://doi.org/10.1109/TVCG.2018.2864903

Sedlmair, M., Meyer, M., & Munzner, T. (2012). Design study methodology: Reflections from the trenches and the stacks. *IEEE Transactions on Visualization and Computer Graphics*, *18*(12), 2431–2440. https://doi.org/10.1109/TVCG.2012.213

Setzer, A. W., & Ferreira, N. J. (2021). O Programa Queimadas do INPE. In A. W. Setzer & N. J. Ferreira (Eds.), *Queimadas e incêndios florestais: Mediante monitoramento orbital* (pp. 43–46). São Paulo: Oficina de Textos.

Shimabukuro, M. H., Flores, E. F., de Oliveira, M. C. F., & Levkowitz, H. (2004). Coordinated views to assist exploration of spatio-temporal data: A case study. In J. C. Robert (Ed.), *Proceedings of the Second International Conference on Coordinated & Multiple Views in Exploratory Visualization*, London, UK (pp. 107–117). Los Alamitos, CA: Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/CMV.2004.1319531

Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings of the 1996 IEEE Symposium on Visual Languages*, Boulder, CO (pp. 336–343). Los Alamitos, CA: Institute of Electrical and Electronics Engineers.

Slattery, P., Saeri, A. K., & Bragge, P. (2020). Research co-design in health: A rapid overview of reviews. *Health Research Policy and Systems*, *18*(1), 17. https://doi.org/10.1186/s12961-020-0528-9

Wijk, J. J. V., & Selow, E. R. V. (1999). Cluster and calendar based visualization of time series data. In G. Wills & D. Keim (Eds.), *Proceedings 1999 IEEE Symposium on Information Visualization (InfoVis'99)*, San Francisco, CA (pp. 4–9). https://doi.org/10.1109/INFVIS.1999.801851

Yang, C.-T., Chen, C.-J., Tsan, Y.-T., Liu, P.-Y., Chan, Y.-W., Chan, W.-C. (2019). An implementation of real-time air quality and influenza-like illness data storage and processing platform. *Computers in Human Behavior*, *100*, 266–274. https://doi.org/10.1016/j.chb.2018.10.009

Zuk, T., & Carpendale, S. (2007). Visualization of uncertainty and reasoning. In A. Butz, B. Fisher, A. Krüger, P. Olivier, & S. Owada (Eds.), *Smart Graphics: 8th International Symposium, SG 2007* (Conference proceedings), Kyoto, Japan (pp. 164–177). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-540-73214-3_15

About the authors

Júlia Rabetti Giannella
juliagiannella@id.uff.br
Universidade Federal Fluminense
Niterói, RJ

Ximena Illarramendi
ximena.rojas@fiocruz.br
Fundação Oswaldo Cruz
Rio de Janeiro, RJ

Renato Campos Mauro
renato.mauro@cefet-rj.br
COPPE/UFRJ e CEFET-RJ
Rio de Janeiro, RJ

Lucas Barcellos Oliveira
lucas.barcellos.oliveira@poli.ufrj.br
Universidade Federal do Rio de Janeiro
Rio de Janeiro, RJ

Igor Falconieri Santiago
igorfalconieri@gmail.com
Universidade Federal do Rio de Janeiro
Rio de Janeiro, RJ

Claudio Esperança
esperanc@cos.ufrj.br
Universidade Federal do Rio de Janeiro
Rio de Janeiro, RJ

Doris Kosminsky
doriskos@eba.ufrj.br
Universidade Federal do Rio de Janeiro
Rio de Janeiro, RJ

Submission date/*Artigo recebido em*: 25/2/2024
Approvement date/*Artigo aprovado em*: 6/9/2024