A Work Project, presented as part of the requirements for the Award of a Master's degree in Business Analytics from the Nova School of Business and Economics.

MAPPING SOCIOECONOMIC ACHIEVEMENT GAPS IN PORTUGAL USING DATA VISUALIZATION

ALICE MARTINS CHAVES

Work project carried out under the supervision of:

Luís Catela Nunes

23/01/2023

Abstract

In this paper, socioeconomic-based indicators of academic achievement and their relationships with municipal-level variables are transferred from the complex scientific research to an interactive and intuitive analysis, with data visualization tools as the primary driver. Three distinct dashboards using Portuguese-specific data were created in Tableau and successfully tested with a diverse range of end-users. These dashboards offer guidance about the different indicators of academic achievement, allow the visualization of regional disparities and temporal tendencies, and show the relationship with several municipal-specific characteristics. Users are now able to analyze and combine multiple characteristics without even scrolling. Dashboards are available at: https://public.tableau.com/app/profile/alice.chaves/viz/DashboardTese-MappingSocialInequalities/DashboardIntrodutorio.

Keywords

Dashboard; Data Visualization; Tableau; Education; Academic Achievement Disparities; Regional Disparities

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209). An appreciation to DGEEC for providing access to anonymized education data.

Table of Contents

1. Introduction 4

2. Portuguese Education System 7

3. Data Collection & Indicators/Dimensions Available 8

3.1. Achievement Indicators 8

3.2. Municipalities Characteristics 9

4. Methodology 10

4.1. Dashboards 11

4.2. Data Visualization Platforms 11

4.2.1. Power BI 12

4.2.2. Tableau 12

4.2.3. Comparison of Power BI and Tableau 13

4.3. Dashboard's Development 14

4.3.1. "Thoughtful Planning" 15

4.3.2. "Information Designed" 16

4.3.3. "Refining Your Dashboard" 17

5. Dashboards: Goals, Choices, Challenges and Solutions 18

5.1. Introduction Dashboard 18

5.2. Map Dashboard 20

5.3. Scatter Plot Dashboard 22

6. User Testing 23

7. Conclusion 24

8. References 26

1. Introduction

Academic achievement disparities between students from high and low socioeconomic (SE) backgrounds, also known as "socioeconomic achievement gaps", have been documented in practically every country (Mullis et al. 2020; OECD 2019). A significant number of factors are suggested as responsible for this situation and various studies have been conducted in different countries to measure the impact of socioeconomic disadvantages on school results (Reardon, Kalogrides, and Shores 2019).

In the US case, this relationship has already been studied and concluded as relevant in the past (Neff 1938). A few years later, Coleman et al. (1966) showed that the primary determinants are related to parental qualifications, wealth, and ethnicity and that school resources were of considerably less importance. Furthermore, children raised in low-income households are more likely to experience severe stress or other clinical issues that might impair cognitive development (Nelson and Sheridan 2011). In addition, higher-income households have access to more stimulating learning environments (Egalite 2016).

Achievement gaps can also be found within the same region/country as a consequence of variables such as SE status, race/ethnicity, gender, and mobility, among others (Ladson-Billings 2006). For instance, the analysis performed in the study by Reardon, Kalogrides, and Shores (2019) mentions that achievement gaps between pupils from various racial/ethnic groups are frequently wide in the United States with these inequalities differing significantly across states. According to this study, racial achievement disparities in school districts and metropolitan regions are strongly correlated with differences in family resources and segregation patterns. These factors, which include numerous racial inequalities in children's opportunities and experiences in their families, communities, and schools, are probably connected to achievement gaps.

Furthermore, there is significant diversity in intergenerational mobility across the United States. For example, in Charlotte, there is a 4.4% chance that a child from a family in the bottom quintile would make it to the top quintile of the national income distribution, whereas there is a 12.9% chance in San Jose (Chetty et al. 2014). The previous author suggests five characteristics highly associated with regional heterogeneity in upward mobility: lower residential segregation, lower income disparity, better primary schools, more social capital, and higher family stability.

For Portugal, the scenario is not different, as it is considered one of the EU nations with the biggest asymmetry in wage income distribution, with qualifications and professions as the main sources of salary differentiation (Rodrigues, Figueiras, and Junqueira 2012). The conclusions of Carneiro (2008) on Portugal are comparable to those of the Coleman study, identifying family background as the most significant factor affecting academic achievement with school resources having a minor influence on student results. This study highlights that better-educated parents can offer their children better living and school settings, which results in higher test scores.

Differences across regions are also found in Portugal. The study “Regional Disparities in Socioeconomic-Based Achievement Measures” (Nunes et al. 2021) suggests that municipalities located in the Norte and Área Metropolitana de Lisboa regions have stronger SE-based achievement gaps. In contrast, the school results of students from lower socioeconomic backgrounds in Área Metropolitana de Lisboa and Alentejo regions are worse in comparison to municipalities in other regions. Additionally, municipalities with worse results for low SE backgrounds students are often areas where the achievement gaps are higher. Furthermore, municipal-level factors such as average hourly salaries, percentage of college-educated employees, Gini Index, abstention rate, crime and unemployment rates, and school segregation are linked to student's academic achievements in these locations.

The majority of the studies mentioned usually use extensive statistical procedures and computations, and their studies can be based on a large number of statistical indicators which add complexity to the results and can make them challenging to analyze. The use of a data visualization platform is considered fundamental in order to better comprehend and explore all the results of a study. These platforms allow the development of dashboards, which can be defined as customizable visual displays that provide the information needed to accomplish one or more goals with simple, clear, and intuitive display techniques (Few 2006).

This approach has already been adopted in other projects. For the US, the Opportunity Atlas (United States Census Bureau et al. n.d.), a result of collaboration between researchers at the United States Census Bureau, Harvard University, and Brown University, provides an interactive mapping tool, enabling users to investigate how children from various backgrounds fare as adults throughout the United States. For Portugal, Fundação Belmiro de Azevedo (n.d.) offers dynamic dashboards to display a set of indicators and explanatory measures, in order to better comprehend the current status, evolution tendencies, and structural dynamics of the Portuguese educational system.

The contribution of this paper is to allow the visualization of the relationships between SE conditions and students' performance at school across different Portuguese regions, by providing easy and intuitive dashboards where users can explore and study this subject without the requirement for a statistical background/knowledge. The indicators included in the dashboards will be the ones obtained by Nunes et al. (2021), mainly the indicators for the SE-based achievement gap, the performance of low SE students, and the average academic scores per region. Municipal-specific indicators are also available to explore.

Three separate dashboards were created using Tableau as the visual analytics platform. The first dashboard serves as an introductory dashboard, where the user can learn about the other two dashboards available and the sorts of analytics they provide. The second dashboard consists of a Mainland Portugal map that allows the user to understand the distribution of educational disparities across different Portuguese areas based on SE-based Achievement indicators. Finally, the third dashboard employs a scatter plot to highlight the relationship between these and municipal-specific characteristics. All these dashboards are available at https://public.tableau.com/app/profile/alice.chaves/viz/DashboardTese-MappingSocialInequalities/DashboardIntrodutorio.

The remaining of this paper is organized as follows. Section 2 of this paper provides a brief overview of the Portuguese educational system. A third section explains the data used and where it was obtained. The fourth section will go over the methodology. A fifth section will go through each dashboard's goals, visualization choices, challenges, and solutions. The sixth section will concentrate on the sharing of the dashboards with users who are unfamiliar with the data and include their feedback. The seventh segment concludes.

2. Portuguese Education System

From the age of six to eighteen, education is mandatory in Portugal. The Portuguese educational system comprises Preschool education and basic, secondary, and higher education (Direção-Geral de Estatísticas da Educação e Ciência (DGEEC) and Ministério da Educação e Ministério da Ciência 2021; Ministério da Educação 2007). Preschool education is optional and intended for children aged between 3 and the age of starting compulsory schooling (Direção-Geral da Educação 2019). Basic education lasts around nine years and is divided into three cycles, each lasting four (from first to fourth grade, or typical ages 6 to 9), two (from fifth to sixth grade, or typical ages 10 to 11), and three years (from 7th until 9th grade, typical ages 12 to 15), respectively (Ministério da Educação 2007). At this point in their schooling, pupils are mostly following the same course of study. It is in secondary school, which lasts three years, that students must choose between two main tracks: academic or professional courses. Academic coursework is primarily targeted at students who want to continue their studies to higher education, whereas professional education focuses on students that want to enter the workforce (OECD 2014).

In Portugal, during the period of analysis, all students had to take national standardized examinations in Mathematics and Portuguese Language in the fourth, sixth, and ninth grades[1]. In the 12th grade, the exams varied according to students' track and course choices (OECD 2020).

3. Data Collection & Indicators/Dimensions Available

As previously stated, the dashboards were developed using the indicators and the data used in the Nunes et al. (2021) study. In the first part of this section, we will focus on the SE Based achievement measures. Secondly, we will describe the metrics for the municipalities' SE characteristics.

3.1. Achievement Indicators

The following sources of data are used in Nunes et al. (2021) to construct the academic achievement indicators.

The students' results on national examinations are provided by *JNE* (Júri Nacional de Exames) and the dataset provides information on exam scores from 2006 to 2018. The *MISI* dataset, controlled by the Portuguese Ministry of Education's DGEEC, provides data on SE characteristics, area of residence, and schools of the students enrolled in Portuguese public education system between 2006 and 2018. Data on the SE characteristics of students enrolled in the Portuguese private school system are imputed using student-level SE data from *PISA*, which is managed by the Organization for Economic Cooperation and Development (*OECD*).

Three main indicators of SE-based academic achievement are available and were chosen due to the different insights each one provides. Firstly, the Performance of Low SE Students indicator corresponds to the average academic achievement of students in the bottom quartile of the national SE index distribution in each region. This measure can be seen as related to

[1] Currently, Portugal uses a dual system for national student assessments: in the 2nd,5th, and 8th grade (mid-cycle), diagnostic testing is administered to students even though they do not have an influence on the final grade and are used mainly to help assess the knowledge acquired in different courses and allow professors to understand where students are facing more difficulties; in the 9th and 12th (end-of-cycle), students sit summative national examinations, in the 9th-grade students perform national exams for Mathematics and Portuguese Language only (OECD 2020).

educational opportunities and is used to analyze whether students coming from lower SE backgrounds perform differently in different regions. Secondly, the SE-Based Achievement Gap corresponds to the difference in average academic achievement between students from the top and bottom quartiles of the national SE index distribution in each region. This indicator allows for measuring the inequalities in achievement between students from high and low socioeconomic backgrounds in different regions. Finally, the Average Results for all Students indicator corresponds to the average academic achievement for all students in each region which allows understanding of which municipalities students have better and worse achievement. This last indicator does not account for the SE status of the students and was selected only for comparison purposes.

Academic achievement was measured using different indicators such as: the proportion of students that achieved a specific grade (9^{th} or 12^{th}) without any previous grade retentions; the average score in the 9^{th} grade national exam; the proportion of students with scores above a passing threshold (4^{th},6^{th}, and 9^{th} grades).

Concerning dimensions, it is possible to explore the results according to different geographic levels (municipalities, NUTS, or National level), different courses (Mathematic, Portuguese, or both), and grades (4^{th}, 6^{th}, 9^{th}, 12^{th} grades). All combinations are available in Appendix Figure 1.

3.2. Municipalities Characteristics

In the second half of this section, concerning the metrics for the SE characteristics of municipalities, the disparities throughout regions are investigated using data from all 278 municipalities in Portugal and covering different dimensions. In terms of labor market data, it was utilized the dataset from *Quadros de Pessoal* which includes information on all non-government employees. Data from the *Comissão Nacional de Eleições* was gathered on abstention rates and voting percentages. To conclude, the data on crime and unemployment rates were collected from two different sources: *Direção Geral da Política de Justiça* and *Instituto do Emprego e Formação Profissional*, respectively.

The second section data, similar to the previous one, added new indicators to analyze for each municipality. These indicators can be divided into five main topics: Labor Market Characteristics, which include the average gaining per hour, the labor market index, the percentage of workers with a college education, the percentage of high-ranked workers and the unemployment rate; Crime, which includes the number of crimes reported in a year per 1000 inhabitants; Income Inequality, such as the Gini Index; Election Outcomes, with the percentage of left-wing political parties and abstention rate; School Segregation, measured between and within schools in terms of academic achievement. All municipal-specific indicators are available in Appendix Figure 2.

Furthermore, it was required to utilize a shapefile obtained from *Portal de Dados Abertos da Administração Pública*, to construct a Portuguese map with each municipality specified since Tableau does not identify Portuguese municipalities by default.

For both sections of the data, a large number of indicators and dimensions were developed, which brings significant complexity to the analysis of the results. This situation demonstrates the importance of developing dashboards to provide the capability of exploring all the different combinations and gaining additional insights on the subject without becoming overwhelmed with all the information and calculations.

4. Methodology

To make the SE-based indicators per region available, it was first necessary to choose the most suitable data visualization platform for this paper’s purpose. Additional research was performed to determine which platforms are in higher ranking positions as well as their benefits and drawbacks when compared. Furthermore, to build the dashboards, the main guide was the steps mentioned by Tableau (n.d.). Additionally, the research performed in the studies of Moody (2009) and Pappas and Whitman (2011) had a high influence on the dashboard’s development.

4.1. Dashboards

The term "dashboard" had its origin in the 19th century to describe a barrier placed in front of a carriage to avoid the horse's hooves from kicking up mud. Later, automobile dashboards were developed to communicate to the driver the status of the vehicle's numerous components (Janes, Sillitti, and Succi 2013).

Nowadays, a dashboard can be defined as “a specific type of page where you can insert reports, graphs, charts, and KPI lists in order to create a central location for functionally relevant information” (Withee 2010). Dashboards are useful in a range of industries and enterprises because of their highly customizable features. By employing visuals like tables, graphs, and charts, others who are less familiar with the data may easily understand the narrative it conveys or the insights it delivers. They are used by professionals to analyze complex data or by experts to track or display data to non-specialists (Tableau n.d.).

Using dashboards as a tool to convey information can be extremely powerful since, as humans, we prefer getting information in visual form and can comprehend it quickly since vision occupies almost a quarter of our brains, which is more than all our other senses combined (Kosslyn 1985).

4.2. Data Visualization Platforms

The Gartner's Magic Quadrant for Analytics and Business Intelligence Platforms 2022 (Appendix Figure 3) mentions Microsoft, Salesforce (Tableau), and Qlik as the three market leaders for the 2022 year (Mitchell 2022). Additionally, concerning Forbes Advisor Ratings (1-5), the best data visualization tools software of 2022 are Microsoft Power BI (4.5), Tableau (4.4), and Qlik Sense (4.4). This rating considers characteristics including cost, customer service, customization, data security, and the capacity to exchange data with other applications (Haan and Watts 2022). In this sub-section, it is provided a brief overview as well as a comparison of the two main platforms highlighted in both researches, Power BI and Tableau.

4.2.1. Power BI

Microsoft describes Power BI as "a collection of software services, apps, and connectors that work together to turn your unrelated sources of data into coherent, visually immersive, and interactive insights". Power BI Desktop, Power BI Service, and Power BI mobile apps are the three core elements of this platform. These elements are created to enable the most efficient development, transmission, and consumption of business insights (Hart et al. 2022).

Gartner's report commended Microsoft Power BI for its synergy with Office 365, Azure Synapse, and Teams, which has contributed to its strong market growth (Richardson et al. 2021). The price/value combination was also mentioned as a factor of differentiation since it offers high service quality with a low pricing strategy. Microsoft highlights other advantages such as security, speed, and constant updates based on recommendations from its members worldwide (Power BI n.d.).

Power BI allows users to connect to data from a wide variety of sources. These available connectors can be divided into four main groups: "File Group", "Database Group", "Azure Group" and "Group Other". The "File Group" has as its main types Excel, CSV, and XML. The "Database Group" mainly consists of SQL databases, Oracle and IBM DB2. The "Azure Group" includes all the Microsoft azure databases and storage. Finally, "Group Other" incorporates programs such as SharePoint list, Google Analytics, and GitHub (fluentpro 2019).

4.2.2. Tableau

According to Tableau's definition, Tableau is a "visual analytics platform transforming the way we use data to solve problems-empowering people and organization to make the most of their data" (Tableau n.d.). Tableau's website emphasizes Tableau Desktop, Server, Cloud, and Prep as its primary offerings. The first one supports offline usage, intuitive drag-and-drop navigation, and data access and analysis. Tableau Server is a web-based platform used to host and manage Tableau data sources and dashboards. Tableau Cloud enables the sharing, automation, and collaboration capability of Tableau Server in a fully cloud-based solution. Tableau Prep is used for shaping and cleaning data. Tableau Public - a free online platform - is available for publicly sharing data visualizations (Tableau n.d.).

Gartner's report mentioned the strength of its user-friendly interface that allows users to visually examine their data easily. The capacity to handle massive quantities of data, the usage of different scripting languages and the active tableau community with a wealth of online resources are other benefits of this platform (Absent Data n.d.).

In terms of data connections, users have access to various data. Tableau groups these into three main categories: "Tableau Server", "To a File" and "To a Server". The first allows to retrieve data sources created and stored in a Tableau Server. Different options such as Excel and CSV files, text files, or JSON files can be found using the connection from a file. The "To a Server" allows connections to SQL Server, Oracle, and Amazon Redshift (Segal 2021).

4.2.3. Comparison of Power BI and Tableau

The previous comparisons of Power BI and Tableau reveal no substantial differences between both tools. However, there are important distinctions to mention. The articles conducted by Biswal (2022) and InterviewBit (2022) suggest the following differences:

- Tableau has higher performance in handling large quantities of data. Additionally, Tableau does not impose row or dimension restrictions or establish a limit on the number of data points that can be displayed.
- Due to Power BI incorporation and similar structure to other Office 365 programs, such as Excel, which many users are familiar with, it is most of the time referred to as easier to learn. As a result, Power BI is utilized by both novice and experienced users, whereas Tableau is mostly used by analysts and advanced users.
- Power BI has been adding more data sources, however, it is still limited to a smaller number when compared to Tableau.
- As a Microsoft product, Power BI is not suitable for Mac users, in contrast to Tableau which works for both Windows and Mac users.

- In contrast to Tableau, which has thousands of active users engaging in the Tableau online community, Power BI is relatively newer to the market. As a result, it has a smaller community.
- The pricing is one of the biggest distinctions between these two systems. Due to Tableau's much greater cost, not many businesses can afford to license it.

After careful consideration of the main differences and considering personal experience with both platforms, the dashboards of this paper were developed using Tableau mainly due to its flexibility, advanced features for data visualization, and its speed regardless of the data size.

4.3. Dashboard's Development

Dashboards are a great tool for data visualization, however, many of them fail to deliver to their full extent (Few 2007). While there is a significant amount of attention and consideration regarding the content provided, the visual form of this content is most of the time taken as non-relevant, focusing more on the aesthetics than on its value (Moody 2009).

To develop the dashboards further explored in this paper was necessary to assure they were effective and usable for the audience with data visualizations that can clearly communicate the data's message, are simple to understand, don't take up too much space, and are appealing. Guidance for the users regarding the next steps is crucial to enable them to achieve their goals and get concrete insights from the data. Placement, attention cues, cognitive burden, and interactivity all contribute significantly to the success of a dashboard and the overall value it gives to its users (Pappas and Whitman 2011).

To guarantee the success of the dashboards, the main steps taken into consideration followed the "10 Best Practices for Building Effective Dashboards" developed by Tableau (n.d.) to better guide its users. This article considers three different factors that make the difference when building a dashboard: "Thoughtful Planning", "Informed Design" and "Refining your Dashboard".

4.3.1. "Thoughtful Planning"

The main steps mentioned by the author in this section are to understand who the end-users of the dashboards are, which display size to use, and the time it takes to load the data.

Regarding the first step, it is crucial that while developing the dashboards we are aware of how knowledgeable the audience is about the data and the topic covered. After getting this information, we can proceed to the decision of the best way to transmit the story behind the data, the most suitable graphics to address the problem, and the level of interactivity that the dashboard should allow its users to perform (Pappas and Whitman 2011). For this paper's purpose, we consider the main audience Portuguese individuals who are interested in the fields of education and social-economic inequalities, and who are familiar with technology but still beginners with data visualization solutions. As a result, the dashboards provide adequate explanations of the indicators displayed, state what each filter allows to accomplish, the message each graphic addresses, and the primary steps that should be performed to reach the dashboard's purpose.

The second step is the display size consideration, emphasizing the importance of understanding where the dashboards will be accessed from. In response to this query, a specific size suitable for the vast majority of desktops available in Tableau was used, as this was considered to be the primary tool used by end users. Additionally, the visualizations were carefully designed and planned to be presented on a display that can be seen all at once, without scrolling, as a dashboard is intended to be viewed fast (Pappas and Whitman 2011; Few 2006).

The third step mentioned is to be aware of the loading times, as it might have a negative impact on its users. To cover this problem, all the data used was carefully transformed and cleaned. Also, since the dashboards are not using real-time data, the "Extract" method was chosen instead of a live data source. On top of that, the latest version of Tableau is always used to optimize performance (Tableau n.d.).

4.3.2. "Information Designed"

The next section of the dashboard's development is the "Informed Designed" which focuses on three key steps: how viewers will "read" the dashboard, the limitation of visualizations and colors, the interactivity, and the dashboard's configuration.

The first step focuses on being cautious regarding how the audience addresses the dashboard as it is necessary to assure the right movement within the dashboard, maintaining a logical structure. People consume information from left to right, and anything placed in the center of the panel must be significant since it will be seen first by the user (Wickens and Hollands 2000). Also, Moody (2009), in his study, advises the use of a "Navigational Map", allowing the user to know where it is, where to go, if it is on the right path and if already accomplished the final goal. Based on these suggestions, all dashboards were created from left to right, specifying the actions that the user should do. Furthermore, the main graphic for each dashboard is in the middle of the panel, capturing the user's immediate attention. In terms of the "Navigational Map", an introductory dashboard (section 5.1) was created to ensure that the user has a broad sense of where to get the information he desires. Furthermore, all dashboards provide the option of navigating back or forward in the dashboards.

The next step refers to the limitation of visualizations, colors, and objects. The author advises building panels with no more than three visualizations and that having too many colors can overload the viewers and slow down their analysis or block it. The color scheme was one of the dashboards' primary areas of concentration since color is one of the visual factors with the greatest cognitive impact (Moody 2009), and color variations are noticed three times more quickly than differences in shape (Lohse 1993). This paper's dashboards followed a consistent color scheme, with no more than three color hues. In addition, when referring to the SE-based indicators the color red was always used across dashboards and the blue color was introduced for the municipality-specific indicators. Color brightness was also considered because bright colors might distract the user (Rasmussen, Chen, and Bansal 2009). Regarding the limitation of graphical items, none of the dashboards display more than two graphics.

The third stage in this section focuses on letting end-users interact with the visualizations, which was one of the most critical and carefully considered features while developing the dashboards. The author highlights the relevance of these interactions in engaging the user and how they might be performed through filter boxes, highlighting, or parameter actions. All the dashboards presented in this paper allow the users to perform multiple interactions based on their preferences. The possibility of filtering the data according to the characteristics chosen allows the display of the information within a smaller space and keeps the attention of the user (Pappas and Whitman 2011). In this paper's dashboards, the user may pick several characteristics to display, such as the indicator, region, grade, and courses, and the visuals will automatically adjust to those selections; also, the user can emphasize any municipality of his choosing using the highlighting box. Furthermore, most of the dashboard titles and legends change dynamically based on the user's selections, assisting the user in understanding his modifications and how to interpret them.

The last step of this section suggests that the dashboard's configuration must follow a largest-to-smallest approach, which was taken into consideration, considering a "Theme", "Workbook", and "Worksheet" order.

4.3.3. "Refining Your Dashboard"

The last section of these practices focuses on tooltips, non-relevant information, and the testing of the dashboards with a group of end-users.

After completing the overall structure and visuals of the dashboards, tooltips are highly suggested by the authors to increase the user's comprehension. Tooltips are created automatically by Tableau but can be fully customized, allowing the allocation of the most significant components first, followed by the least important. Furthermore, the incorporation of graphs in the tooltips enables the user to investigate deeper into the data without occupying extra space in the dashboard. In all the dashboards of this paper, the tooltips were explored and developed to facilitate the user's understanding of the data, emphasizing the relevant information to absorb. Moreover, the dashboard map (section 5.2) contains a tooltip that provides additional information on the evolution of the SE-based indicator per region. This allowed the offering of extra details and information towards a more specific component of the data (Pappas and Whitman 2011).

The second part of this section highlights that everything presented must have a purpose. Removing items deemed unneeded in certain graphs, such as a title or axis name, as well as simplifying the colors and structure, improves the user's understanding of the panel. The dashboards presented in this study were designed to be as simple as possible. Further, all the considered non-relevant objects were removed as it was a priority to include only what was absolutely needed for the end goal, keeping the display easy to read and understand (Few 2007).

Finally, the author emphasized the importance of assessing the effectiveness of the dashboards with possible end-users and using their feedback to improve them. Two main methods were used to analyze the usefulness of the dashboards produced. The first was a casual encounter, with the purpose of visualizing the exploration sequence followed by the participants. The second method was in a virtual environment and consisted of a tutorial video explaining the dashboard's goal and main functionalities. In the end, the users were questioned about the main issues found and possible changes/improvements (further explained in section 6).

5. Dashboards: Goals, Choices, Challenges and Solutions

This section will explore in detail the dashboards created, mainly their primary goals, options, difficulties, and solutions.

5.1. Introduction Dashboard

The first dashboard available when accessing the Tableau Public link is an introductory dashboard. Although there is no visualization or data displayed, this dashboard was designed to guide viewers in understanding what additional analysis and visuals they will discover and how to most benefit from them (Figure 1).

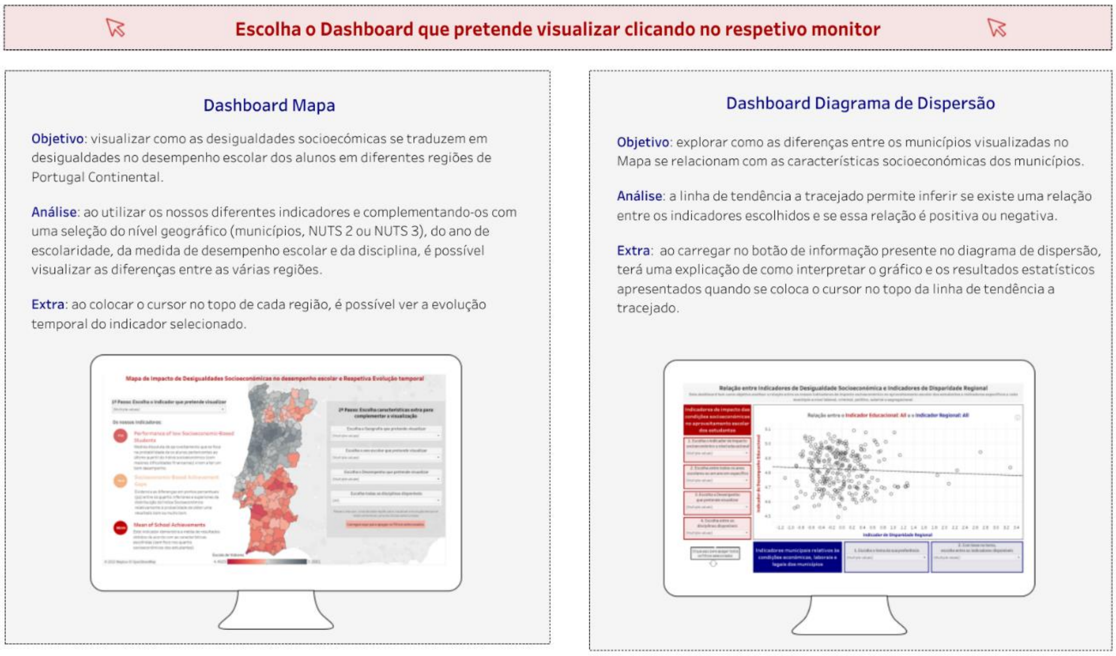

Figure 1- Introduction Dashboard

To keep a simple but concise visualization, the display was divided into two main boxes, each one containing information regarding one of the dashboards. This information involves the goals, analytical capabilities, and other elements that may be used to support the analysis for each of the dashboards. Additionally, to facilitate the user's understanding of how to proceed from this dashboard to the others, it was followed the "Principle of Semantic Transparency" which consists in using a visual representation that also illustrates its meaning (Moody 2009). As a result, images of monitors with the dashboards already shown were employed to provide a depiction of the user's reality when exploring the platform; by clicking on the monitor, the user is directed to the corresponding dashboard.

To finalize, as discussed in the previous section, this dashboard serves as a "Navigational Map" indicating where the user is and where he may go. This is useful since two separate dashboards are available, with different analyses, and the user may become lost if no direction was provided.

5.2. Map Dashboard

The goal of this dashboard is to visualize how socioeconomic inequalities translate into academic achievement disparities for students across different regions of Portugal (Figure 2).

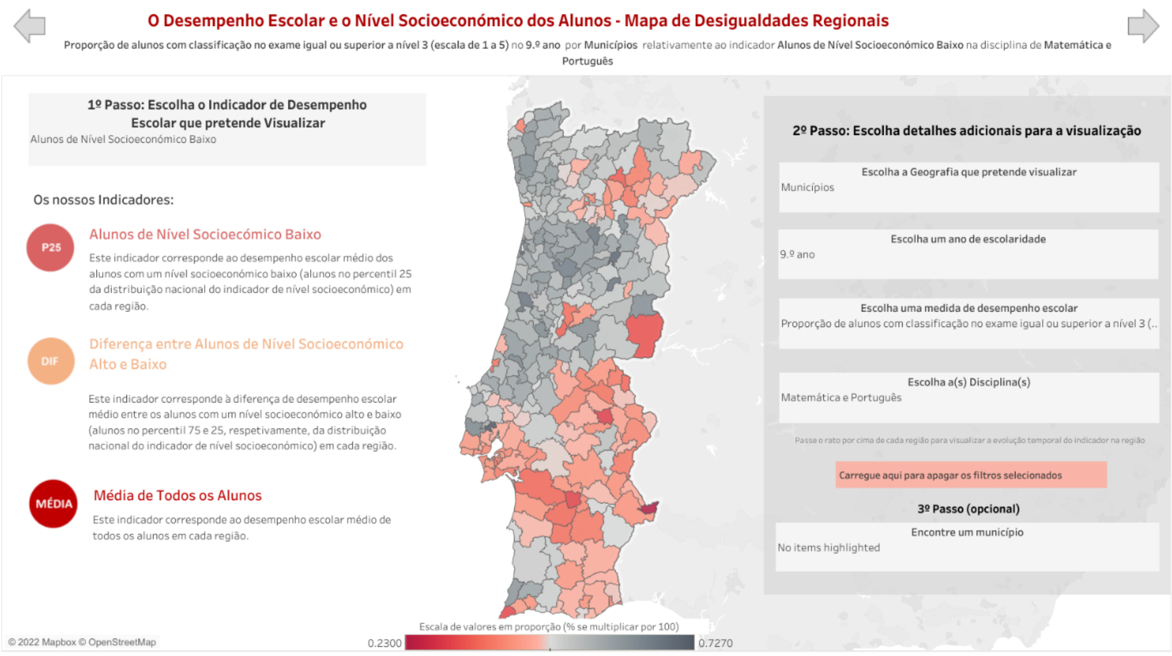

Figure 2 - Map Dashboard

A map graphic was chosen to address this objective since geographic representation can express complex data such as location characteristics (MacEachren and Kraak 2001) and is also easy to interpret for non-technical users (Osagie et al. 2017).

In terms of analysis, the user can choose the indicator to investigate further and complement it with additional dimensions such as geographic level, school grade, achievement metric, and, if applicable, a specific course. In addition, a supplemental visual is accessible in the map graphic's tooltip to visualize the historical evolution of the attributes chosen over time.

Several challenges had to be overcome during the development of this dashboard to ensure its effectiveness. To begin with, and as previously stated, the structure and presentation of the dashboard are crucial for its success, thus it was essential to keep the design simple and straightforward, making sure the user was not overwhelmed with information. One solution to this problem was to add the temporal evolution as the map tooltip, so that the user still has access to the analysis, but it does not take up extra space in the dashboard. In addition, just two colors were utilized, and only the information deemed required was presented.

Secondly, this dashboard involves significant user engagement to select which indicators and characteristics to display, making it critical to ensure that a suitable direction toward the end goal is provided. For that, the dashboard clearly states the steps that the user should do and the respective order. In addition, each filter in the dashboard specifies the action it performs - all actions are made through "selection" - and what aspect of the data will this filter be focused on - indicator, grade, course, or achievement measure.

Furthermore, as noted before, the statistical indicators may be difficult to interpret without any support. As a result, on the left side of the dashboard, a brief description for each of the indicators is provided, using simple and plain terminology, in the same order as they appear in the respective filter.

Additionally, it was necessary to only allow the display of the map graphic once all filters have been selected, having only one data point per region. To solve this issue, a filter was created and applied to the visualization, requiring that the values counted in the data are equal to one.

This dashboard involves multiple filter selections and therefore, the final result might be challenging to interpret immediately. For that reason, as the user chooses the characteristics of the visualization, a dynamic legend will adapt to those choices, developing an easy interpretation of what the user can take from the map. Additionally, a color ramp legend with a small description of the scale used is available. Moreover, removing each filter individually for a new selection might be time-consuming, therefore, to remedy this problem, a reset button was added, making it possible to delete all selections at the same time.

Finally, the geographic level of the municipality separates the map into extremely small sections, making it difficult for the user to locate the place he wishes to investigate further. As a response, a highlighting box with the option to write the name of the municipality or scroll through the list of existing municipalities was added.

5.3. Scatter Plot Dashboard

The last dashboard of this paper aims to explore how the differences between the municipalities visualized in the Map Dashboard are related to their SE characteristics. To achieve this goal, a dynamic scatter plot was developed (Figure 3) since it is an adequate visual to display the relationship between two numeric variables, allowing one to understand the type of relationship, the data points used, and the identification of outliers (Sainani 2016).

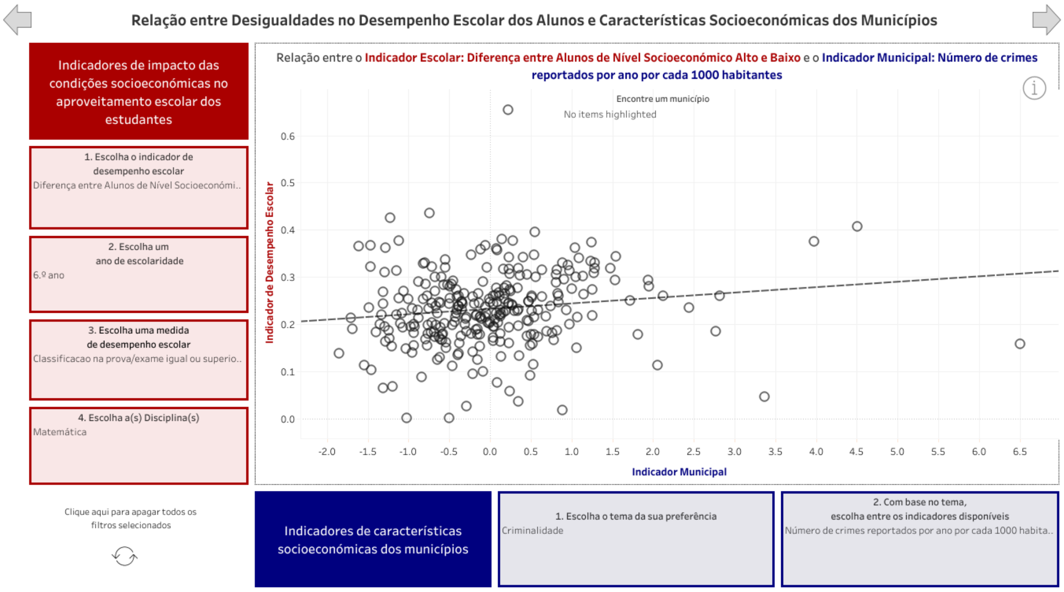

Figure 3- Scatter plot Dashboard

On the vertical axis, we have the SE-based indicators and the average scores. On the horizontal axis, the municipality-specific indicators previously mentioned are added.

When it comes to the analysis, the tendency line makes it simple for the user to determine whether there is a relationship between the selected indicators and whether that relationship is positive or negative. The user has the same selections to perform as in the previous dashboard (except for the region level, which is not available in this dashboard), as well as filters based on municipality features.

Certain challenges had to be overcome when developing the dashboard. Firstly, it was crucial that the user could intuitively understand which filters are impacting which axis. For that, we relied on the dashboard's design, for both the colors and structure. For the school-based indicators, the color red was maintained from the previous dashboard and the blue color was introduced for the municipality-specific data. Therefore, the vertical axis title is colored red and the horizontal one is blue. Additionally, the scatter plot has a dynamic title that varies according to the user's selection, mentioning which indicators are selected using their specific color (blue or red). This challenge is also tackled by the structure followed, as the vertical filters affect the vertical axis, and the horizontal filters influence the horizontal axis.

Secondly, some users might not be familiar with scatter plot analysis. Therefore, an information button was added to the dashboard, identifying how to interpret the graphic based on the tendency line, with a brief explanation of its main statistical elements. Similarly to the map dashboard, it was important to enable the user to easily identify municipalities of his choice without having to search through all the data points. Therefore, a highlighting filter with the municipalities' names was included. Additionally, a significant number of filter selections is also necessary, and consequently, the reset button is available.

6. User Testing

As mentioned in the methodology (section 4), user testing was conducted to measure the dashboard's effectiveness with people non-familiar with the subject or the data. It consisted of two different methods.

The first one was completed in-person, in an informal environment, and individually with the user. The main objectives were to understand how the user "reads" the dashboard and the direction followed both inside and across the various dashboards. The majority of the participants easily understood the possibilities of each dashboard and followed the expected order of interaction – introduction dashboard, map dashboard, and lastly, the scatter plot dashboard – and the expected order of reading – from left to right. This method was conducted with six people -

with backgrounds ranging from nursing, engineering, education, sports, economics, and human resources – and took on average seven minutes per participant.

The second method was conducted virtually using a video tutorial on how to analyze and explore the dashboards. This video involves a begin-to-end analysis of the dashboards. During the video, each dashboard was presented, focusing on the purpose, type of analysis, and extra features available. This approach was conducted with eleven participants, with a variety of backgrounds such as: education, management, data analytics, and finance. The average time to complete this experiment was around twelve minutes, with seven minutes reserved to visualize the video and five minutes to explore the dashboards.

After conducting each method, the participants were asked to provide feedback on their experience, highlighting features they found more useful and possible ways to improve the analysis. Based on this feedback, the following characteristics were added: a direct search of the municipality name, a dynamic legend to help the result's interpretation, the rephrasing of less clear vocabulary, and the increasing of important information size. The overall feedback was positive for both methods, with users eager to explore the dashboards and learn more about the subject. No major difficulties were pointed out and therefore the dashboards were finalized.

7. Conclusion

Various studies throughout the years have suggested different factors that can be related to students' inequalities in academic achievement. However, most of them are not easy to interpret for non-mathematically driven minds. With Portugal as the focus of the analysis, an interactive approach that would allow visualizing the disparities in academic achievements across regions, as well as which factors are more related to those disadvantages, was missing. To address these issues, the SE-based indicators on school achievement, the average scores, and municipality-specific indicators were taken from the current study of Nunes et al. (2021) and used to develop a non-technical analysis of this subject.

Three alternative dashboards were developed using the aforementioned indicators. The choice to use this approach was mainly due to the dashboard's effectiveness in sharing information in a simple and interactive manner (Tableau n.d.). As for the data visualization platform chosen to develop these dashboards, Tableau was selected due to its flexibility, wide variety of visualization features, and capability of handling large volumes of data. Therefore, the main steps followed while building these dashboards were based on Tableau (n.d.).

The initial dashboard does not offer any data analysis or visualization; instead, it gives the user a broad summary of what is accessible in the other dashboards, acting as a jumping-off point and source of assistance. The second dashboard is focused on the three main indicators, namely, the performance of Low SE students, the Socioeconomic Achievement Gap, and Average Academic Scores. This dashboard allows the visualization of the disparities across regions for each one of these indicators, based on multiple dimensions such as geographic level, grade, academic achievement measure, or course. Additionally, this dashboard provides the possibility for the user to visualize a temporal trend of how this indicator has been behaving throughout the years. Finally, the third dashboard seeks to address the relationship between the three indicators mentioned and municipality-specific indicators; for that reason, a dynamic scatter plot was built.

After a careful construction of these visuals, user testing was conducted to assess their effectiveness with potential end-users. For that, two major approaches were employed: casual in-person meetings and a tutorial video on how to make the most out of the dashboards. The participants' feedback was taken into consideration, and improvements were made accordingly.

In future work, it would be important to include other characteristics in the dashboards, such as the gender and ethnicity of the students in order to measure their impact on academic achievement. Additionally, there is still a wide range of visualizations available to explore that can provide further insights regarding the subject.

8. References

Absent Data. n.d. "Advantages and Disadvantages of Tableau." Absent Data. Accessed December 10, 2022. https://absentdata.com/advantages-and-disadvantages-of-tableau/.

Biswal, Avijeet. 2022. "Power BI Vs Tableau: Difference and Comparison." November 29, 2022. https://www.simplilearn.com/tutorials/power-bi-tutorial/power-bi-vs-tableau.

Carneiro, Pedro. 2008. "Equality of Opportunity and Educational Achievement in Portugal." *Portuguese Economic Journal* 7 (1): 17–41. https://doi.org/10.1007/s10258-007-0023-z.

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfield, and Robert L. York. 1966. "Equality of Educational Opportunity." Washington.

Direção-Geral da Educação. 2019. "Educação de Infância." October 3, 2019. https://www.dge.mec.pt/perguntas-frequentes-faq.

Direção-Geral de Estatísticas da Educação e Ciência (DGEEC), and Tecnologia e Ensino Superior Ministério da Educação e Ministério da Ciência. 2021. *Educação e Formação Em Portugal*.

Egalite, Anna. 2016. "How Family Background Influences Student Achievement: Can Schools Narrow the Gap?" *Education Next* 16 (2): 70–78.

Few, Stephen. 2006. *Information Dashboard Design: The Effective Visual Communication of Data*. Beijing: Cambridge [MA]: O'Reilly.

———. 2007. "Why Most Dashboards Fail." www.PerceptualEdge.com.

fluentpro. 2019. "Power BI Data Sources: What You Can Connect & How." Fluentpro. December 23, 2019. https://fluentpro.com/blog/power-bi-data-sources-what-you-can-connect-how/.

Fundação Belmiro de Azevedo. n.d. "EDUSTAT - Observatório Da Educação." Accessed December 11, 2022. https://www.edustat.pt/.

Haan, Kathy, and Rob Watts. 2022. "The Best Data Visualization Tools of 2022." Forbes Advisor. October 18, 2022. https://www.forbes.com/advisor/business/software/best-data-visualization-tools/.

Hart, Michele, Michael Blythe, Alex Buck, Mona Berdugo, Kesem Sharabi, Maggie Sparkman, and Kent Sharkey. 2022. "What Is Power BI?" Power BI. January 28, 2022. https://learn.microsoft.com/en-us/power-bi/fundamentals/power-bi-overview.

InterviewBit. 2022. "Power BI vs Tableau: Full Comparison [2022]." InterviewBit. July 1, 2022. https://www.interviewbit.com/blog/power-bi-vs-tableau/.

Janes, Andrea, Alberto Sillitti, and Giancarlo Succi. 2013. "Effective Dashboard Design." *Cutter IT Journal* 26 (1): 17–24.

Kosslyn, Stephen M. 1985. "Graphics and Human Information Processing: A Review of Five Books." *Journal of the American Statistical Association* 80 (391): 499–512.

Ladson-Billings, Gloria. 2006. "From the Achievement Gap to the Education Debt: Understanding Achievement In." *Educational Researcher* 35 (7): 3–12. https://www.jstor.org/stable/3876731.

Lohse, Gerald Lee. 1993. "A Cognitive Model for Understanding Graphical Perception." *Human–Computer Interaction* 8 (4): 353–88. https://doi.org/10.1207/s15327051hci0804_3.

MacEachren, Alan M., and Menno-Jan Kraak. 2001. "Research Challenges in Geovisualization." *Cartography and Geographic Information Science* 28 (1): 3–12. https://doi.org/10.1559/152304001782173970.

Ministério da Educação. 2007. *Educação e Formação Em Portugal*. Edited by Ministério da Educação.

Mitchell, Charlie. 2022. "Gartner Magic Quadrant for Analytics and Business Intelligence Platforms 2022." CX Today. March 30, 2022. https://www.cxtoday.com/data-analytics/gartner-magic-quadrant-for-analytics-and-business-intelligence-platforms-2022/.

Moody, Daniel. 2009. "The Physics of Notations: Toward a Scientific Basis for Constructing Visual Notations in Software Engineering." *IEEE Transactions on Software Engineering* 35 (6): 756–79. https://doi.org/10.1109/TSE.2009.67.

Mullis, Ina V.S., Michael O. Martin, Pierre Foy, Dana L. Kelly, and Bethany Fishbein. 2020. *TIMSS 2019 International Results in Mathematics and Science*. TIMSS & PIRLS International Study Center, Lynch School of Education and Human Development, Boston College and International Association for the Evaluation of Educational Achievement (IEA). https://timssandpirls.bc.edu/timss2019/international-results/.

Neff, W. S. 1938. "Socioeconomic Status and Intelligence: A Critical Survey." *Psychological Bulletin* 35 (10): 727–57. https://doi.org/10.1037/h0055707.

Nelson, Charles, and Margaret Sheridan. 2011. "Lessons from Neuroscience Research for Understanding

Causal Links between Family and Neighbourhood Characteristics and Educational Outcomes." In *Whither Opportunity*, edited by Greg Duncan and Richard Murnane, 27–46. New York: Russell Sage Foundation.

Nunes, Miguel, Luís Catela Nunes, Pedro Carneiro, and Pedro Freitas. 2021. "Regional Disparities in SocioeconomicBased Achievement Measures."

OECD. 2014. "Education Policy Outlook Portugal." *Education Policy Outlook*. Paris.

———. 2019. "Students' Socio-Economic Status and Performance." In *PISA 2018 Results (Volume II): Where All Students Can Succeed*, 2:49–62. PISA. Paris: OECD Publishing. https://doi.org/10.1787/b5fd1b8f-en.

———. 2020. "Education Policy Outlook in Portugal." *OECD Education Policy Perspectives*. Paris. https://doi.org/https://doi.org/10.1787/0e254ee5-en.

Osagie, Edobor, Mohammad Waqar, Samuel Adebayo, Arkadiusz Stasiewicz, Lukasz Porwol, and Adegboyega Ojo. 2017. "Usability Evaluation of an Open Data Platform." In *ACM International Conference Proceeding Series*, Part F128275:495–504. Association for Computing Machinery. https://doi.org/10.1145/3085228.3085315.

Pappas, Lisa, and Lisa Whitman. 2011. "Riding the Technology Wave: Effective Dashboard Data Visualization." *LNCS*. Vol. 6771. Cary.

Power BI. n.d. "Why Microsoft Power BI." Accessed December 9, 2022. https://powerbi.microsoft.com/en-us/why-power-bi/.

Rasmussen, Nils, Claire Y. Chen, and Manish Bansal. 2009. *Business Dashboards: A Visual Catalog for Design and Deployment*. Hoboken: John Wiley & Sons, Inc.

Reardon, Sean F, Demetra Kalogrides, and Kenneth Shores. 2019. "The Geography of Racial/Ethnic Test Score Gaps." *American Journal of Sociology* 124 (4): 1164–1221. https://doi.org/http://doi.org/10.1086/700678.

Richardson, James, Kurt Schlegel, Rita Sallam, Austin Kronz, and Julian Sun. 2021. "Gartner Magic Quadrant for Analytics and Business Intelligence Platforms." Gartner. February 15, 2021. https://www.gartner.com/en/documents/3996944.

Rodrigues, Carlos Farinha, Rita Figueiras, and Vitor Junqueira. 2012. "Desigualdade Económica Em

Portugal.” Lisboa.

Sainani, Kristin L. 2016. “The Value of Scatter Plots.” *PM&R Journal* 8 (12): 1213–17. https://doi.org/10.1016/j.pmrj.2016.10.018.

Segal, Mark. 2021. “Connecting Tableau to Different Types of Data Sources.” Zuar. July 29, 2021. https://www.zuar.com/blog/connecting-tableau-to-data-sources/.

Tableau. n.d. “10 Best Practices for Building Effective Dashboards.”

———. n.d. “Optimize Workbook Performance.” Tableau. Accessed December 13, 2022a. https://help.tableau.com/current/pro/desktop/en-us/performance_tips.htm.

———. n.d. “Our Products.” Accessed December 9, 2022b. https://www.tableau.com/products.

———. n.d. “What Is a Dashboard? A Complete Overview.” Tableau. Accessed December 9, 2022c. https://www.tableau.com/learn/articles/dashboards/what-is.

———. n.d. “What Is Tableau?” Accessed December 9, 2022d. https://www.tableau.com/why-tableau/what-is-tableau.

United States Census Bureau, Nathan Hendren, Brown Friedman, and Raj Chetty. n.d. “Opportunity Atlas.” Accessed December 11, 2022. https://www.opportunityatlas.org/.

Wickens, Christopher D., and Justin G. Hollands. 2000. *Engineering Psychology and Human Performance*. 3rd ed. Prentice Hall Inc.

Withee, Ken. 2010. *Microsoft Business Intelligence for Dummies*. Indianapolis: Wiley Publishing.

9. Appendix Figures

Appendix Figure 1 - Possible combinations to explore in the dashboards using the different indicators and dimensions available

Possible Combinations of Dimensions

				Anos				
Indicador	Desempenho	Grade	Disciplina	g1	g2	g3	g4	gall
Alunos de Nível Socioeconómico Baixo	Classificação média em pontos (escala de 0 a 100)	9.º ano	Matemática e Portu..	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
			Matemática	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
			Português	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
	Proporção de alunos com classificação no exame igual ou superior a nível 3 (escala de 1 a 5)	4.º ano	Matemática e Portu..	2008-2010	2011-2012	2013-2015		2008-2015
			Matemática	2008-2010	2011-2012	2013-2015		2008-2015
			Português	2008-2010	2011-2012	2013-2015		2008-2015
		6.º ano	Matemática e Portu..	2008-2010	2011-2012	2013-2015		2008-2015
			Matemática	2008-2010	2011-2012	2013-2015		2008-2015
			Português	2008-2010	2011-2012	2013-2015		2008-2015
		9.º ano	Matemática e Portu..	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
			Matemática	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
			Português	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
	Proporção de alunos com classificação no exame igual ou superior a nível 4 (escala de 1 a 5)	4.º ano	Matemática e Portu..	2008-2010	2011-2012	2013-2015		2008-2015
			Matemática	2008-2010	2011-2012	2013-2015		2008-2015
			Português	2008-2010	2011-2012	2013-2015		2008-2015
		6.º ano	Matemática e Portu..	2008-2010	2011-2012	2013-2015		2008-2015
			Matemática	2008-2010	2011-2012	2013-2015		2008-2015
			Português	2008-2010	2011-2012	2013-2015		2008-2015
		9.º ano	Matemática e Portu..	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
			Matemática	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
			Português	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
	Proporção de alunos..	9.º ano	Matemática e Portu..	1996-1998	1999-2000	2001-2002		1996-2002
	Proporção de alunos..	12.º ano	Matemática e Portu..	1993-1995	1996-1997	1998-1999		1993-1999
	Proporção de alunos..	12.º ano	Matemática e Portu..	1993-1995	1996-1997			1993-1997
	Proporção de alunos..	12.º ano	Matemática e Portu..	1993-1995	1996-1997			1993-1997
Diferença entre Alunos de Nível Socioeconómico Alto e Baixo	Classificação média em pontos (escala de 0 a 100)	9.º ano	Matemática e Portu..	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
			Matemática	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
			Português	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
	Proporção de alunos com classificação no exame igual ou superior a nível 3 (escala de 1 a 5)	4.º ano	Matemática e Portu..	2008-2010	2011-2012	2013-2015		2008-2015
			Matemática	2008-2010	2011-2012	2013-2015		2008-2015
			Português	2008-2010	2011-2012	2013-2015		2008-2015
		6.º ano	Matemática e Portu..	2008-2010	2011-2012	2013-2015		2008-2015
			Matemática	2008-2010	2011-2012	2013-2015		2008-2015
			Português	2008-2010	2011-2012	2013-2015		2008-2015
		9.º ano	Matemática e Portu..	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
			Matemática	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
			Português	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
	Proporção de alunos com classificação no exame igual ou superior a nível 4 (escala de 1 a 5)	4.º ano	Matemática e Portu..	2008-2010	2011-2012	2013-2015		2008-2015
			Matemática	2008-2010	2011-2012	2013-2015		2008-2015
			Português	2008-2010	2011-2012	2013-2015		2008-2015
		6.º ano	Matemática e Portu..	2008-2010	2011-2012	2013-2015		2008-2015
			Matemática	2008-2010	2011-2012	2013-2015		2008-2015
			Português	2008-2010	2011-2012	2013-2015		2008-2015
		9.º ano	Matemática e Portu..	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
			Matemática	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
			Português	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
	Proporção de alunos..	9.º ano	Matemática e Portu..	1996-1998	1999-2000	2001-2002		1996-2002
	Proporção de alunos..	12.º ano	Matemática e Portu..	1993-1995	1996-1997	1998-1999		1993-1999
	Proporção de alunos..	12.º ano	Matemática e Portu..	1993-1995	1996-1997			1993-1997
	Proporção de alunos..	12.º ano	Matemática e Portu..	1993-1995	1996-1997			1993-1997
Média de Todos os Alunos	Classificação média em pontos (escala de 0 a 100)	9.º ano	Matemática e Portu..	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
			Matemática	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
			Português	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
	Proporção de alunos com classificação no exame igual ou superior a nível 3 (escala de 1 a 5)	4.º ano	Matemática e Portu..	2008-2010	2011-2012	2013-2015		2008-2015
			Matemática	2008-2010	2011-2012	2013-2015		2008-2015
			Português	2008-2010	2011-2012	2013-2015		2008-2015
		6.º ano	Matemática e Portu..	2008-2010	2011-2012	2013-2015		2008-2015
			Matemática	2008-2010	2011-2012	2013-2015		2008-2015
			Português	2008-2010	2011-2012	2013-2015		2008-2015
		9.º ano	Matemática e Portu..	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
			Matemática	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
			Português	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
	Proporção de alunos com classificação no exame igual ou superior a nível 4 (escala de 1 a 5)	4.º ano	Matemática e Portu..	2008-2010	2011-2012	2013-2015		2008-2015
			Matemática	2008-2010	2011-2012	2013-2015		2008-2015
			Português	2008-2010	2011-2012	2013-2015		2008-2015
		6.º ano	Matemática e Portu..	2008-2010	2011-2012	2013-2015		2008-2015
			Matemática	2008-2010	2011-2012	2013-2015		2008-2015
			Português	2008-2010	2011-2012	2013-2015		2008-2015
		9.º ano	Matemática e Portu..	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
			Matemática	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
			Português	2008-2010	2011-2012	2013-2015	2016-2018	2008-2018
	Proporção de alunos..	9.º ano	Matemática e Portu..	1996-1998	1999-2000	2001-2002		1996-2002
	Proporção de alunos..	12.º ano	Matemática e Portu..	1993-1995	1996-1997	1998-1999		1993-1999
	Proporção de alunos..	12.º ano	Matemática e Portu..	1993-1995	1996-1997			1993-1997
	Proporção de alunos..	12.º ano	Matemática e Portu..	1993-1995	1996-1997			1993-1997

Anos Intervalo broken down by Anos vs. Indicador, Desempenho, Grade and Disciplina. The data is filtered on Geografia, which keeps Municípios.

Appendix Figure 2 - All the municipal-specific indicators available in the Scatter Plot Dashboard

Regional Indicator Topic	Regional Indicator Name
Labor Market Characteristics	The average earnings per Hour
Labor Market Characteristics	The labor market index
Labor Market Characteristics	The percentage of college-educated workers
Labor Market Characteristics	The percentage of high-ranked workers
Labor Market Characteristics	The unemployment rate
Crime	The number of crimes reported in a year per 1000 inhabitants
Income Inequality	The Gini Index
Election Outcomes	The percentage of left-wing political parties
Election Outcomes	The abstention rate
School Segregation	School segregation within the same school in the 4th grade
School Segregation	School segregation within the same school in the 6th grade
School Segregation	School segregation within the same school in the 9th grade
School Segregation	School segregation between schools in the 4th grade
School Segregation	School segregation between schools in the 6th grade
School Segregation	School segregation between schools in the 9th grade

Appendix Figure 3 - Magic Quadrant for Analytics and Business Intelligence Platforms

Source: Gartner (March 2022)