DASHBOARD FOR TERRITORIAL DATA ANALYSIS: POTENTIAL, SDMX AND INTEGRATION WITH OTHER SOURCES

Andrea Arru, Paolo Misso

Abstract. Dashboards have become an indispensable tool for the effective and timely representation and analysis of socio-economic phenomena. Organizations that deal with statistics increasingly promote the use of this tool, managing to satisfy the needs and interests of an increasingly wider audience of users. This work presents an effective approach to analyse multi-source territorial data on tourism, using dashboards of the most popular business analytics platforms in order to evaluate their potential. One of the most relevant features of dashboards is the ability to visualize and analyse data up to municipal and monthly levels, offering a temporal and territorial granularity particularly suited for in-depth studies. Furthermore, the platforms allow users to cross-reference tourist information with other territorial variables, such as demographic, economic or infrastructural data, enriching the contextual analysis. The systems are updated both using the ISTAT REST APIs through the SDMX protocol, ensuring dynamic and continuously updated access to official data, and datasets already available previously downloaded and structured.

1. Introduction

Dashboards have become a key tool for data management and analysis, capable of effectively supporting evidence-based decision-making processes. Thanks to their ability to aggregate, summarise and visualize complex information in an intuitive manner, dashboards are widely used across both public and private sectors to monitor phenomena, assess performance, and communicate results to various economic stakeholders (Eurostat, 2021, OECD, 2021).

Beyond their technical and analytical functions, dashboards can also be framed within broader debates on *data governance* and *public sector innovation*. In fact, the capacity to integrate heterogeneous sources and ensure data quality resonates with the literature on data governance frameworks (Khatri & Brown, 2010; Janssen, Charalabidis & Zuiderwijk, 2012). Similarly, the emphasis on openness, transparency and accountability directly connects dashboards to research on open government data and public communication of statistics (Bertot, Jaeger & Grimes, 2010; Meijer, 2015).

An increasingly significant dimension in the development of dashboards is the geographical one, that is, the spatial representation of data—particularly in reference to official statistics produced by public institutions, statistical agencies and governmental bodies (Istat, 2025). The use of thematic maps, Geographic Information Systems (GIS), and spatial visualisations enables the contextualization of information in relation to territory, making local dynamics, territorial disparities, critical issues and latent opportunities immediately visible. This form of representation enhances the accessibility and interpretability of official data for policymakers, citizens and practitioners (Paul A. Longley *et al.,* 2015) and has been successfully applied in geospatial dashboards for monitoring complex urban and territorial systems (Jing *et al.,* 2019, Schaffert *et al*., 2024), but also aligns with the stream of research on *e-governance dashboards* as instruments for decision support and smart governance in the public sector (Scholl & Scholl, 2014).

A key feature of modern dashboards is their ability to integrate heterogeneous data sources — combining real-time and historical information from censuses, registries, monitoring systems, IoT streams, and more (Stehle & Kitchin, 2019). Interoperability among different sources and formats enables the construction of rich, real-time updateable and highly customisable information frameworks.

The main objective of this work is to develop an effective tool for the preliminary exploration of complex phenomena, characterised by low economic investment, rapid implementation times, ease of deployment, and continuous update capabilities.

A specific focus has been placed on the data ETL process (Extract, Transform, Load), with particular attention to the use of the SDMX format (Statistical Data and Metadata eXchange), which makes it possible to leverage pre-normalised data structures and explicit metadata to streamline the automation of information flows.

2. Data

For the implementation of the multi-source dashboard model, the tourism sector was selected, as it is particularly well suited to showcasing the integration of data from both official and complementary sources. Tourism is, in fact, a complex phenomenon characterised by highly heterogeneous spatial and temporal dynamics, and effective analysis requires the combination of statistical, geographical and economic information.

In developing the dashboard, official data from various Istat information systems (such as IstatData, Demo, Aster, etc.) were integrated. These systems collect multi-source data derived from statistical surveys and administrative records. The datasets cover a wide range of dimensions and are typically available in standard formats such as CSV, SDMX and Excel.

This approach highlights the potential of multi-source dashboards in providing a comprehensive view of tourism-related phenomena, supporting analyses of territorial distribution, seasonality, economic impacts, and accommodation capacity, with dynamic updates and intuitive geospatial visualisations. In this study, multi-decade time series were used, which required a careful preliminary phase of data verification, harmonisation and reconstruction, given that the original datasets were often incomplete or inconsistent.

2.1. Dataset

The datasets used to populate the dashboard are largely derived from census surveys, while a smaller portion originates from Istat's processing of administrative data sources. A selection of these datasets is listed below:

- Annual population movement and calculation of resident population
- Reconstruction of the demographic balance (2001–2018)
- Accommodation capacity of hospitality establishments
- Client movement in hospitality establishments
- ASIA Local Units Register
- Istat processing of open data on tax declarations (MEF – Department of Finance)
- Maritime transport
- Air transport
- Boundaries of administrative units for statistical purposes.

The data were downloaded in the various available formats, primarily CSV and SDMX, in order to compare them in terms of download speed, ease of processing, and update capabilities. CSV is the most widely used format, whereas the IstatData and Aster platforms also provide data in SDMX format. In some cases, Excel format is still available; however, it requires more complex transformation procedures to enable automated integration into the ETL workflow.

2.2. SDMX

SDMX has been analysed in greater depth to better assess its potential and limitations in data downloading and uploading. Data extraction from the IstatData platform was carried out through interaction with the SDMX endpoint of Istat (https://esploradati.istat.it/SDMXWS/), implemented by means of scripts in R with the aim of automating access.

The Single Exit Point (SEP) is based on an API that follows the REST architecture and allows the sending of HTTP requests, structured in the form of queries (EUROSTAT 2021, SDMX community 2024). Within the SDMX framework, queries are mainly distinguished into two categories:

- Structural metadata query:
  https://ws-entry-point/**resource**/**agencyID**/**resourceID**/**version**/**itemID**
  ?**queryStringParameters**
- Data query:
  https://ws-entry-point/**resource**/**flowRef**/**key/providerRef**
  ?queryStringParameters

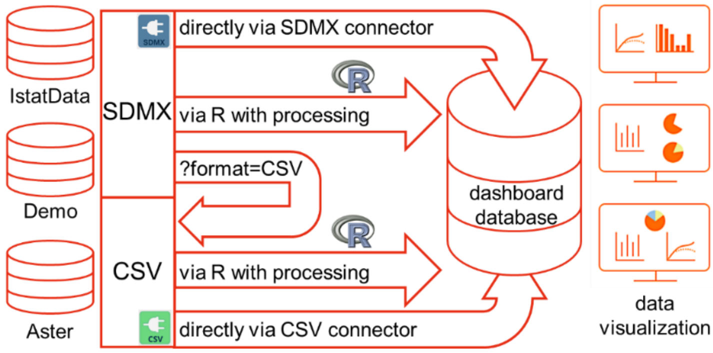
These queries allow the user to specify, for example, the variable to be analysed, the territorial level, the time range, and other filters. However, a limitation that still remains is the lack of clear and easily understandable documentation: the syntactic rules to be followed are not always explained in detail, which often makes the manual construction of queries a complex and error-prone task. To facilitate the use of the SEP, query builders are provided to guide users in composing the query string, making the tool more accessible even to less experienced users.

With the aim of comparing different access methods, the following three solutions were tested:

- direct use of an HTTP client, such as a browser, which allows simple and immediate access;
- use of R, through the RSDMX library;
- use of connectors, that is, tools enabling direct data integration with visualisation applications such as Power BI, MicroStrategy and Tableau.

From the perspective of integration within an ETL automation process, these methods exhibit very different characteristics. Using a browser is suitable for exploratory purposes or occasional access but cannot be automated efficiently. In contrast, employing R with RSDMX allows for the writing of repeatable, programmable, and schedulable scripts, making it suitable for automated pipelines. Finally, connectors can offer a good level of automation if integrated into tools that support scheduled data refreshes, but they tend to be less flexible during intermediate transformation steps compared to scripting languages such as R or Python.

Figure 1 – *Dashboard data feeding scheme via direct connector or processing in R.*



The SDMX format includes, alongside the data, a detailed metadata structure organised according to the Data Structure Definition (DSD), which formally describes the dimensions, measures, attributes, and associated value domains. The DSD is essential for understanding how to correctly read and query an SDMX dataset: each dimension corresponds to a statistical variable, while the measures represent the observed values and the attributes specify their context or quality.

To efficiently construct the tables to be loaded into the dashboard, the opportunity was taken to use codelists, that is, sets of valid codes associated with a given dimension or attribute. Since these are often shared across multiple datasets, codelists tend to be broad and generic. The use of tools capable of navigating and filtering the metadata enabled automation of the process, selecting only the combinations that are actually valid between codes and dimensions.

Within the ETL process actually implemented, access to metadata proved to be a crucial step: it allowed the construction of dynamic and automatable queries, the avoidance of interpretation errors, and the efficient application of filters upstream of extraction. However, this phase required the use of tools capable of correctly reading and interpreting the DSD. In particular, specialised R packages such as rsdmx, readsdmx, and RJSDMX were employed.

2.3. Reconstruction of Territorial Time Series

To ensure the temporal consistency of the territorial time series, a data harmonisation procedure was applied to municipal data by aligning them to the current administrative boundaries. This approach helps mitigate the distorting effects arising from territorial changes that occurred during the observation period, such as mergers, splits, creations, cessations, or alterations of municipal boundaries.

The reconstruction is based on a methodology inspired by that used for the creation of the 8milaCensus information system (ISTAT, 2015), although in this work exclusive reference was made to information concerning the creation and cessation of municipalities, without considering minor boundary changes. The temporal assignment of data was carried out through the reconstruction of historical municipal statistical codes, using information obtainable from SITUAS (SITUAS, 2024; ISTAT, 2024).

The historical series of municipal data for the period 2013–2023 was reconstructed using the administrative geography in force as of 1 January 2023 as a reference. In particular, the total number of Italian municipalities decreased from 8,092 in 2013 to 7,901 in 2023, following mergers and the establishment of new entities. Among the latter are the municipalities of Mappano (TO), established in 2017, and Misiliscemi (TP), established in 2021, which do not have complete historical series over the entire considered interval. For these municipalities, data prior to their respective establishment are not directly available and were not subject to disaggregated reconstruction; therefore, their series are partial, starting from the years of their actual creation.

Aligning to the 2023 geography also involved reassigning historical data in correspondence with municipal mergers, according to aggregation criteria consistent with the founding acts and territorial continuity.

3. Dashboard

Among the software solutions for creating analytical dashboards, Tableau, Microsoft Power BI, MicroStrategy, and ArcGIS Dashboards have emerged as the most widely adopted internationally. The choice of Tableau for this project is motivated by specific technical features that make it particularly well-suited to handling and visualising data from multiple sources. In particular, Tableau offers broad compatibility with heterogeneous data sources, enabling rapid integration and updating of datasets, and it incorporates advanced GIS functionalities that allow for the creation of interactive maps, spatial analysis, and thematic overlays without the need for external tools. The platform also stands out for the quality of its visualisations, native interactivity, and the ability to build exploratory analysis pathways through dynamic filters and geographical drill-downs. Tableau is supported by a large user community and proves to be an effective tool for dashboard creation, suitable for both expert users and those with limited technical skills. (Salesforce, 2023).

This section does not present a data analysis, but rather an exploration of the key components of the dashboard dedicated to the tourism sector, with the aim of illustrating its main features and potential. The data displayed provides information at municipal level, allowing for a detailed exploration of the territory. The intention is to provide an overview of the visual and functional components, highlighting how these can support the identification of patterns, trends, and anomalies during consultation.

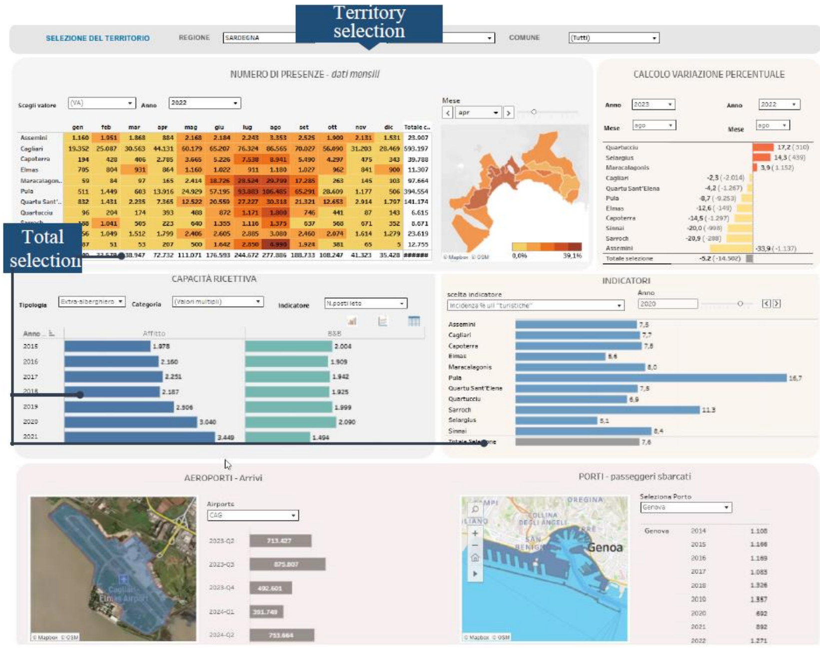
The tool is structured to integrate heterogeneous information sources – including data on tourist flows, accommodation capacity, and other relevant variables – presented through customisable and scalable graphic visualisations. The multi-level territorial filter forms the core element that governs the entire interface: the user can guide the analysis by progressively selecting region, province, and municipality, building flexible exploration paths focused on the area of interest.

The prototype employs a combination of interactive visualizations, including thematic maps for spatial distribution, time-series charts to track historical trends, and comparative dashboards to highlight differences across municipalities. These choices follow established principles in data visualization and statistical communication (Few, 2013; Kelleher & Wagener, 2011; Börner *et al.*, 2019), aiming to make complex administrative and demographic data both accessible and interpretable.

A distinctive feature of the dashboard is the modularity of the spatial component, which allows for contextual geographical selection: users can query data either at a granular level (single municipality) or in aggregate form, assembling customised sets of territories. This configuration enables comparative and cumulative analyses on different geographical aggregates – such as homogeneous areas, functional zones, or sub-regional contexts – even when these do not coincide with administrative boundaries.

Through the functionalities of the dashboard, the spatial, temporal, and structural dimensions that describe the tourism phenomenon can be analysed in detail, allowing users to grasp its complexity and the interrelationships between the variables involved. In particular, for data with a strong temporal component – such as monthly tourist overnight stays – a heatmap has been implemented, enabling users to quickly identify peaks in tourism pressure throughout the year, revealing recurring patterns or seasonal anomalies. This visualisation is integrated with an interactive geographical map that spatially locates the information.

Figure 2 – *Overview of the Dashboard.*



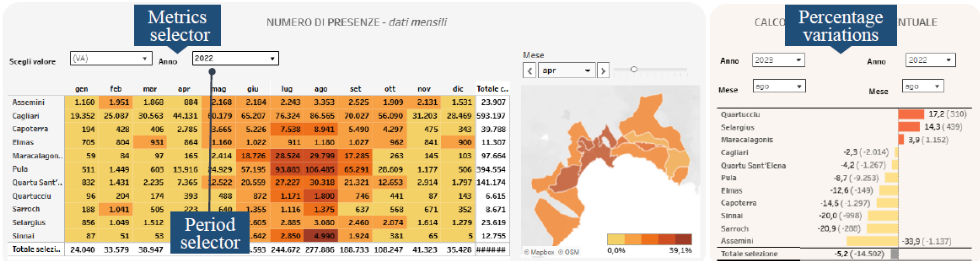
The interface allows for territorial selection through dynamic filters (at regional, provincial, and municipal levels) and displays key indicators for each individual municipality or for the entire selected area.

Thanks to the use of Tableau parameters, the dashboard offers a flexible and customisable configuration of analytical elements. In particular, a function has been integrated to calculate percentage variations, allowing users to compare the same indicator across different periods, based on the selected area. This option proves especially useful for analyses aimed at assessing the temporal evolution of the phenomenon, such as comparing consecutive tourist seasons or evaluating the impact of local policy interventions.

The dashboard allows users to choose the type of metric to be displayed, switching between absolute values, normalised values, or percentage changes, depending on the type of comparison or interpretation intended.

The tool allows the same data to be viewed in different ways, both through charts and tables, in a dynamic manner. The interface enables users to easily switch from a thematic map view to sortable interactive tables, or to bar charts and line charts, while keeping the applied territorial and temporal filters unchanged.

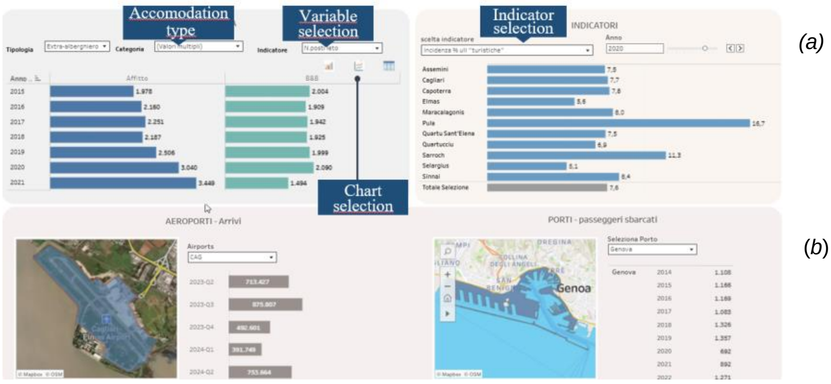
Figure 3 – *Top section of the dashboard with territorial selectors.*



Monthly data on tourist arrivals are presented in a table and a thematic map. The dashboard enables users to select the metric and the period of analysis, and provides a customisable tool for calculating percentage changes.

An additional filter allows the selection of indicators of interest, enabling the analysis to focus on specific aspects of the tourism phenomenon, such as the incidence of foreign visitors, accommodation capacity by type, or occupancy rates. All these indicators are calculated directly within Tableau.

Figure 4 – *Section of the dashboard dedicated to dimensions and indicators (a). Infrastructure data on ports and airports (b).*



On the left: modular selection via cascading filters for the type of accommodation facility, and selection buttons to activate the different graphical representations. On the right: selection and display of indicators. On the bottom: example of a query on georeferenced data, integrating official data on passenger traffic at airports and ports

Tableau integrates GIS capabilities that can be easily connected to the data within the dashboard's database, enabling the dynamic display of indicators associated with the selected geographical entities. This approach proves particularly useful, for example, when analysing airport or port infrastructures in relation to tourist access flows.

4. Conclusions

In the present work, an interactive dashboard has been developed in Tableau, which can also be used as a personal tool, with the aim of integrating, enriching, and expanding the functionalities offered by publicly available statistical data warehouses. Although these public querying systems represent a useful resource for accessing official data, they often present limitations in flexible territorial management and in conducting multidimensional analyses. The proposed dashboard, on the other hand, adopts a territorially focused approach, enabling comparative exploration of data across different administrative levels (municipalities, provinces, regions) and facilitating cross-sectional analysis between variables and territories.

Thanks to Tableau's advanced features, users can customise graphical representations, apply dynamic filters, construct composite indicators, and visualise results through thematic maps, interactive charts, and comparative matrices, as well as refresh the dashboard with continuously updated new data.

During the course of the work, the importance of the SDMX format emerged as a standard for building automated ETL procedures, due to its ability to integrate data and metadata within a single structure and the possibility of querying data providers directly through parametric queries. However, some practical drawbacks were also identified, particularly at the initial stage: SDMX files are often 10 to 20 times larger than their CSV counterparts, resulting in longer download times; moreover, reading and interpreting metadata requires specific tools and expertise, at least during the parameterisation phase of the procedures. For these reasons, the CSV format can still represent a valid — and sometimes the only — alternative, especially in cases where providers, as in the case of the Single Exit Point (SEP) of Istat, also allow the automatic export of data already formatted in CSV, thus simplifying integration into ETL workflows.

Tableau offers considerable ease in importing data from various sources; however, it currently lacks a native connector for the SDMX format, which necessitates the development of a custom connector using the appropriate SDK. This represents a limitation that could be overcome in future software versions, hopefully more open to direct support of this standard. Furthermore, although Tableau allows for various data processing and transformation operations, these functionalities are not comparable to the flexibility and power of analytical environments such as R, which remains indispensable for certain advanced transformations.

Finally, a significant aspect of this work concerned the reconstruction of territorial historical series, necessary to ensure data consistency over time in light of frequent changes in administrative geography, particularly at the municipal level. This operation cannot be fully automated due to the heterogeneity of sources and structural variations in historical datasets, which change over time in terms of format, coverage, and variable structure. In this specific case, the reconstruction was carried out partially, limited to situations manageable with the information available in SITUAS, namely cases of extinction and complete acquisition of municipalities by other entities, which still represent the majority of territorial modifications. Despite its limitations, this intervention significantly improved the longitudinal analysis of phenomena at the municipal scale, providing a firmer basis for consistent temporal and territorial comparisons.

References

BERTOT J. C., JAEGER P. T., GRIMES J. M.. 2010. *Using ICTs to create a culture of transparency: E-government and social media as openness and anti-corruption tools for societies*. Government Information Quarterly.

BÖRNER K., BUECKLE A., GINDA M.. 2019. Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. *Proceedings of the National Academy of Sciences*, Vol.116, No. 6, pp. 1857–1864.

EUROSTAT. 2021. *Eurostat Communication and Dissemination Strategy 2021–2024*, https://ec.europa.eu/eurostat/about-us/policies/communication-dissemination.

EUROSTAT. 2021. *Eurostat SDMX RESTful API*, https://ec.europa.eu/eurostat/api/dissemination/swagger-ui.

FEW, S., 2013. *Information Dashboard Design: Displaying Data for At-a-Glance Monitoring*. Analytics Press.

ISTAT. 2015. 8milaCensus: l*a ricostruzione della popolazione legale dei comuni con variazioni nel periodo 1951-2011*, https://ottomilacensus.istat.it/fileadmin/download/Metodologia_ricostruzione_popolazione_legale_1951_2011.pdf

ISTAT. 2023. *Il turismo culturale in Italia: analisi territoriale integrata dei dati*. https://www.istat.it/wp-content/uploads/2023/10/Il-turismo-culturale-in-Italia_Ebook.pdf

ISTAT. 2024. L*'informazione statistica microterritoriale: l'esperienza del registro statistico di base dei luoghi*. https://www.istat.it/wp-content/uploads/2024/12/informazione-statistica-microterritoriale-Ebook.pdf

ISTAT. 2025. *Statistica ufficiale ed esigenze informative del territorio*. https://www.istat.it/wp-content/uploads/2025/03/Statistica-ufficiale-esigenze-informative-del-territorio-Ebook.pdf .

JANSSEN M., CHARALABIDIS Y., ZUIDERWIJK A..2012. Benefits, adoption barriers and myths of open data and open government. *Information Systems Management.*

JING C., DU M., LI S., LI, S..2019. Geospatial Dashboards for Monitoring Smart City Performance. *Sustainability*, Vol. 11, No. 20, 5648. https://doi.org/10.3390/su11205648

KELLEHER C., WAGENER T.. 2011. Ten guidelines for effective data visualization in scientific publications. *Environmental Modelling & Software*, Vol.26, No. 6, pp. 822–827.

KHATRI V., BROWN C. V.. 2010. *Designing data governance*. Communications of the ACM.

MEIJER A., 2015. *E-governance innovation: Barriers and strategies.* Government Information Quarterly.

OECD. 2021. Enhancing the Use of Data for Public Integrity: A Roadmap for OECD Countries. OECD Publishing.

OECD. 2021. *Report on Public Communication: The Global Context and the Way Forward*, OECD Publishing, Paris, https://doi.org/10.1787/22f8031c-en

PAUL A. LONGLEY, MICHAEL F. GOODCHILD, DAVID J. MAGUIRE, DAVID W. RHIND. 2015. *Geographic Information Science and Systems, 4th Edition.* Wiley

SALESFORCE. 2023. Welcome to Tableau Public. A free platform to explore, create, and publicly share data visualizations online, https://public.tableau.com/app/discover.

SCHOLL H. J., SCHOLL M. C.. 2014. Smart governance: A roadmap for research and practice. *iConference 2014 Proceedings*.

SDMX community. 2024. *SDMX 3.0 web services cheat sheet*, *v2.2.0*, https://github.com/sdmx-twg/sdmx-rest/blob/master/doc/rest_cheat_sheet.pdf?raw=true.

SCHAFFERT M., WARCH D., MÜLLE, H. 2024. A Planning Support System for Monitoring Aging Neighborhoods in Germany. *Geomatics*, Vol.. 4, No. 1, pp. 66-80. https://doi.org/10.3390/geomatics4010004

SITUAS. 2024. *Sistema Informativo Territoriale delle Unità Amministrative e Statistiche*, https://situas.istat.it.

STEHLE S., KITCHIN R. 2019. Real-time and archival data visualisation techniquesin city dashboards. *International Journal of Geographical Information Science*, Vol. 34, No. 2, pp. 344–366. https://doi.org/10.1080/13658816.2019.1594823

Andrea ARRU, Istat, andrea.arru@istat.it
Paolo MISSO, Istat, misso@istat.it