

# An Interactive Dashboard for Visualizing Big Spatial-Temporal Data in an Urban Area

Ningchuan Xiao<sup>1</sup>, Shaun Fontanella<sup>1</sup>, Harvey J. Miller<sup>1,2</sup>, Matthew Adair<sup>2</sup>, and Jerry Mount<sup>2</sup>

<sup>1</sup>Department of Geography, The Ohio State University, Columbus, OH, USA

<sup>2</sup>Center for Urban and Regional Analysis, The Ohio State University, OH, USA

\*Email: xiao.37@osu.edu

## Abstract

An interactive dashboard is developed to visualize public and open data sources for an urban area. We formally describe the data sets and the tasks of data query. We then discuss the technical details of the dashboard development.

**Keywords:** Spatial-temporal data visualization, big data, dashboard, databases

## 1 Introduction

Recent years have seen a huge amount of data from a broad range of sources converge at urban areas around the world. In a typical midwest urban setting in the United States, for example, the Central Ohio area has seen the convolutions of environment, transportation, and socio-economic data. These data sets, coming from both traditional and new online sources, often arrive in huge volume and fast velocity; they also have diverse resolutions, purposes, and methods of collection (Chen et al., 2014). They signify a new era of big data (Kitchin, 2014) and present tremendous challenges to policy makers, the general public, and researchers. One prominent challenge is: how to present the data to users so that so they are able to not only find facts from the data, but to obtain trends and relationships of phenomena that are represented by the data?

The purpose of this paper is to discuss a dashboard approach to visualizing big spatial and temporal data. We see visualization as a means of enabling a user to explore the data through interactive graphical tools. As the literature suggests, interactive visual tools are essential to facilitate human perception of the patterns and relationships in the otherwise complicated data sets (Andrienko and Andrienko, 2006; Wagemans et al., 2012). We use the dashboard as a metaphor where a set of graphical panels are designed to show different aspects of the data. Collectively the dashboard provides a holistic view of urban status and dynamics.

The idea of using a dashboard to provide information in graphical and other forms has been increasingly recognized in the big data literature (Kitchin et al., 2015; Lee et al., 2015). This approach has

also been applied by many municipal agencies for cities such as Amsterdam<sup>1</sup>, Dublin<sup>2</sup>, London<sup>3</sup>, Sydney<sup>4</sup>, and states like Iowa<sup>5</sup> to help their citizens better understand their urban environment. However, current dashboards appear to be more focused on the presentation of data (as data sets) and some of them merely serve as a collage of maps. These implementations generally lack exploratory tools that will enable the user to query and seek relationships and trends within and, more importantly, between different phenomena. We aim to address these limitations by designing a highly interactive environment for users to explore the data.

## 2 Methodology

We start our development from a formal understanding of the data and an analysis of the kinds of queries (questions) on the data. We then use recent developments in databases and web-based systems to implement the dashboard.

### 2.1 Data and tasks

We define spatial and temporal data as a set of  $\{u : x \in X, u \in U\}$ , where  $u$  is the measure of a phenomenon in the domain of  $U$  obtained in its context  $x$  in a context domain  $X$ . The context of a data measure is often a combination of spatial and temporal references that have different spatial and temporal scales. Spatial and temporal data sets often exhibit a functional dependency (Ramakrishnan and Gehrke, 2003) between the data values and their context, allowing us to use the context as the unique identifier for the value. This is similar to the functional view of data in the literature (Andrienko and Andrienko, 2006).

The user queries spatial and temporal data sets to address questions for different purposes. These questions are called tasks (Bertin, 2010). For all these queries, the user, explicitly or implicitly, confines the query with a context constraint. The user may query the value, the context, or both, and we call them query targets. Each query may also have its conditions, which can be value, context, or both. For example, the user may want to see the current air quality index in a city, or the trend of air quality in the past 24 hours or longer period of time. We distinguish two kinds of tasks: singular tasks that are related to a single phenomenon and multiplex tasks that always involve data of two or more phenomena or data sets. Table 1 summarizes the kinds of tasks with respect to the multiplicity of phenomena and query targets. We use  $\{u : x \in X, u \in U\}$  and  $\{v : y \in Y, v \in V\}$  to denote two data sets with their values ( $u$  and  $v$ ) and contexts ( $x$  and  $y$ ), respectively.

Tasks listed in Table 1 often require minimal analysis on the data other than retrieving the raw data from the databases. Sometimes the user may be interested in questions that cannot be directly answered by simply retrieving the data values: “are people who tweeted about food concentrated in the same area as those tweeted about sports”? Tasks like this fall into the category of synoptic tasks (Andrienko and Andrienko, 2006). Some spatial (or nonspatial) analysis must be carried out

---

<sup>1</sup><http://citydashboard.waag.org>

<sup>2</sup><http://www.dublindashboard.ie/pages/index>

<sup>3</sup><http://citydashboard.org/london/>

<sup>4</sup><https://cityfutures.be.unsw.edu.au/cityviz/>

<sup>5</sup><https://www.icalliances.org/iowa-dashboard/>

Table 1: Types of tasks for data query.

Type	Query targets	Examples
Singular	$u$	What is the highest air quality index value in the past 24 hours in Columbus?
	$x$	where is the gas station with the highest price now in the city?
	$u, x$	When was the last time we have gas prices lower than \$2 and where?
Multiplex	$u, v$	In the past 24 hours, which demographic group is more exposed to high health risk due to air quality?
	$x, y$	Where are the gas stations with low prices located with respect to the percent of black or other demographic groups?
	$u, x, v, y$	What is the highest air quality index value in the past 24 hours, where it is recorded, and what is the minority population around that station?

in order to answer such questions. For example, a kernel density estimation will be needed to understand the concentration of people who tweeted about food.

## 2.2 Design and Implementation

The above analysis of data and tasks is critical in specifying the requirements of the dashboard. It also provides guidance for two important dimensions of the design and implementation of the system: database and visualization tools.

We developed our dashboard as a web-based visualization platform. On the server side, we collect data for air quality index of PM 2.5, trip length in minutes, bike sharing station, tweets, gas prices from online sources, and demographic data from the U.S. Census Bureau. The online data are “harvested” using server side Python scripts and APIs. These data are then stored in the PostgreSQL databases on the server, where spatial and temporal indexes are used to expedite database query. We built server side PHP programs to retrieve data and to transfer data to the client side in JSON format. On the client side, visualization tools are built using JavaScript and two major libraries: (1) Leaflet<sup>6</sup> for organizing base maps and custom map layers from the data sets, and (2) D3.js<sup>7</sup> for showing data with different plotting techniques. The data shown on the dashboard are subsets for the recent time period (e.g., past 24 hours), while we also provide links to show data for longer time periods.

Three kinds of interactive features are implemented for the graphic panels on the dashboard. Within panel interactivity allows the user to examine the trend while retrieving individual data elements (querying  $u$  or  $x$  from  $\{u : x\}$ ). Interactivity between a plot panel and the map panel enables the user to examine the value and its spatial-temporal contexts (querying  $u, x$  from  $\{u : x\}$ ). Finally interactivity between two or more plot panels and the map panel helps explore multiple phenomena and their contexts (querying  $u, x, v, y$ , or any combinations of them).

<sup>6</sup><http://leafletjs.com>

<sup>7</sup><https://d3js.org>

### 3 Result

We design the dashboard as the front page of a web site called CURIO (Columbus Urban & Regional Information Observatory)<sup>8</sup>. Figure 1 shows the overall interface of the dashboard in a web browser. The top of the screen is a summary computed using the data about transportation, environment, and social aspects of the city. The largest panel shows the base map, along with maps made from other data sets. Figure 1 also shows examples of interactivity within a plot panel (COGO Bikes): when the user moves the cursor through the bar chart, the plot shows the name of the bike sharing station, which is also highlighted as the green dot on the map.

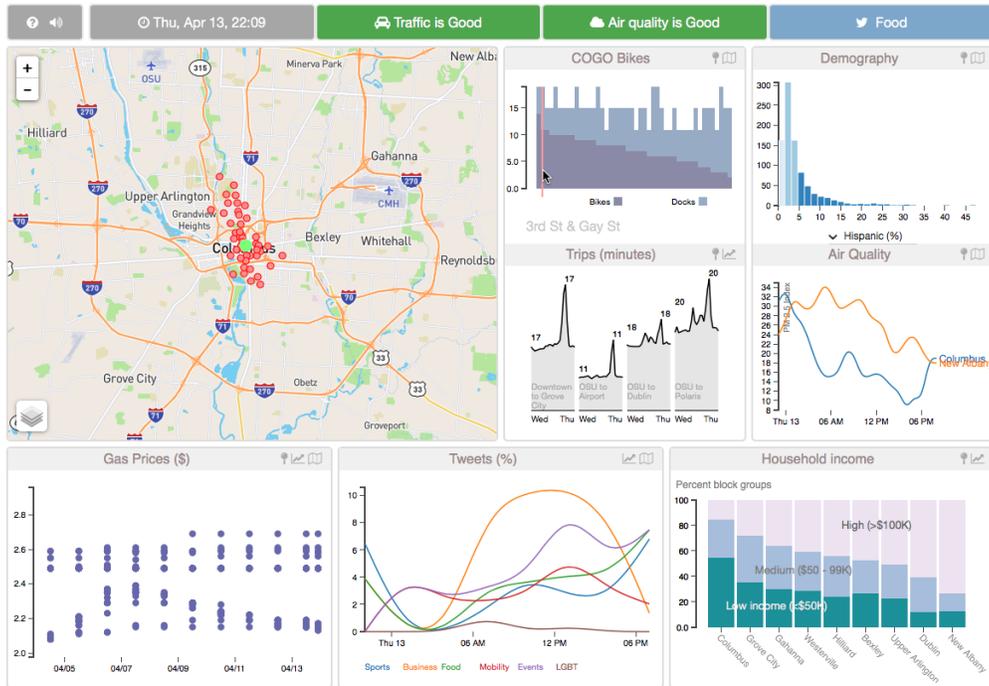


Figure 1: The front page of the dashboard.

Figure 2 shows more interactive features of the dashboard for both singular and multiplex queries. The user can view the overall trend (distribution) of a demographic group (upper-right) while searching for individual values by moving the cursor over the bars. By showing the air quality data on the same map, the user can explore the demographic information in the neighborhoods of the two stations. The user can also highlight a subset of the demographic data (light red areas on the map) and explore their relationship with the air quality sites.

Each plot panel has two or three buttons on the upper-right corner that allow the user to pin the data layer to the map, see the longer term trend of the data, and open a new page with a larger version of that panel's data plotted on a map. By clicking on the map button in the Tweets panel, for example, the user will be able to explore more details about tweets in the region (Figure 3) for both spatial and temporal dimensions.

<sup>8</sup><http://curio.osu.edu>. A development site is at <http://gis.osu.edu/misc/dashboard2>.

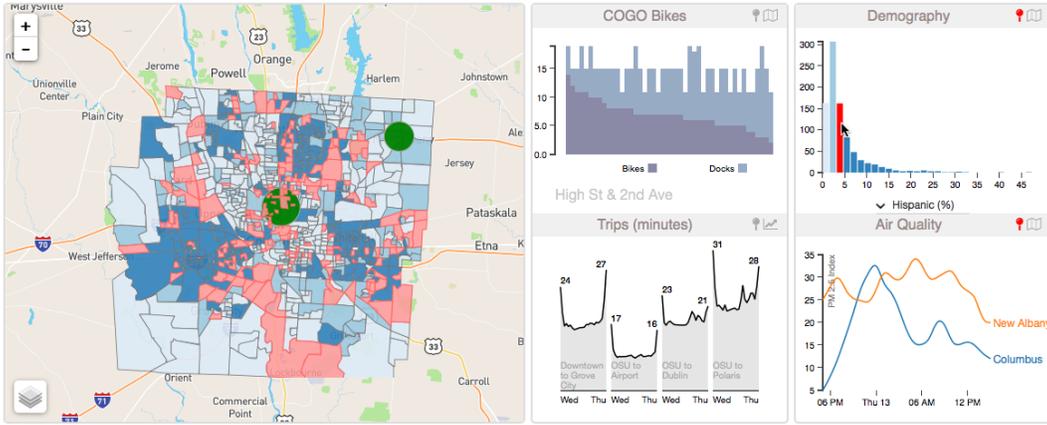


Figure 2: Interactivity within and between panels.

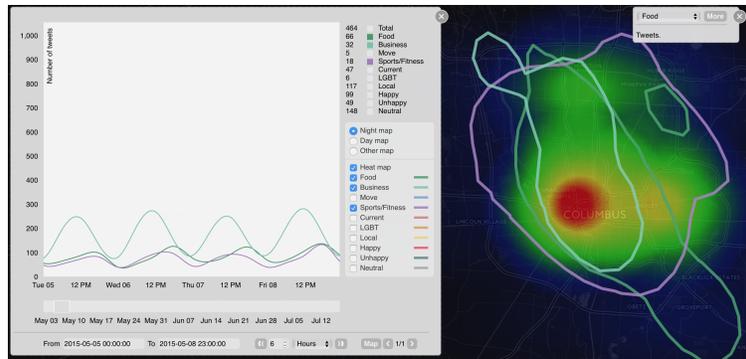


Figure 3: Visualizing tweets in Central Ohio.

## 4 Discussion and Conclusions

The interactive dashboard discussed in this paper can be generalized for other data sets and other areas. We plan to include more visualization techniques and interactive tools in the future. A usability study will also help improve the dashboard.

## 5 Acknowledgements

This research is partly funded by the College of Social and Behavioral Sciences at the Ohio State University.

## 6 References

- Andrienko, N. and G. Andrienko (2006). *Exploratory analysis of spatial and temporal data: a systematic approach*. Berlin: Springer.
- Bertin, J. (2010). *Semiology of Graphics: Diagrams Networks Maps*. Redlands, CA: ESRI Press. First published in French in 1967.

- Chen, M., S. Mao, and Y. Liu (2014). Big data: a survey. *Mobile Networks and Applications* 19(2), 171–209.
- Kitchin, R. (2014). The real-time city? big data and smart urbanism. *GeoJournal* 79(1), 1–14.
- Kitchin, R., T. Lauriault, and G. McArdle (2015). Knowing and governing cities through urban indicators, city benchmarking, and real-time dashboards. *Regional Studies, Regional Science, Journal of the Regional Studies Association* (1), 6–28.
- Lee, D., J. R. A. Felix, S. He, D. Offenhuber, and C. Ratti (2015). Cityeye: Real-time visual dashboard for managing urban services and citizen feedback loops. In *Proceedings of The 14th International Conference on Computers in Urban Planning and Urban Management*.
- Ramakrishnan, R. and J. Gehrke (2003). *Database Management Systems* (3rd ed.). New York: McGraw-Hill.
- Wagemans, J., J. H. Elder, M. Kubovy, S. E. Palmer, M. A. Peterson, M. Singh, and R. von der Heydt (2012). A century of gestalt psychology in visual perception: I. perceptual grouping and figure–ground organization. *Psychological Bulletin* 138(6), 1172.