Exploring Rural Shrink Smart Through Guided Discovery Dashboards

Denise Bradford
University of Nebraska - Lincoln
Lincoln, Nebraska
denise.bradford@huskers.unl.edu

Susan VanderPlas
University of Nebraska - Lincoln
Lincoln, Nebraska
susan.vanderplas@unl.edu

Abstract

Many small and rural places are shrinking. Interactive dashboards are the most common use cases for data visualization and context for exploratory data tools. In our paper, we will explore the specific scope of how dashboards are used in small and rural area to empower novice analysts to make data-driven decisions. Our framework will suggest a number of research directions to better support small and rural places from shrinking using an interactive dashboard design, implementation and use for the every day analyst.

Keywords: Interactive Dashboards, Exploratory Data Analysis (EDA), Guided Discovery

1 Research Problem

With the amount of publicly open-source data, a proliferation of visualization dashboards has increased in nearly every industry (**?**). A dashboard in its fundamental form, a dashboard supports a way of presenting and making sense of complex data to better enable and support decision making.

Some communities continue to thrive as they lose population because they adapt, maintaining quality of life and community services for residents while investing in the future. This process, *smart shrinkage*, is important for rural areas who have experienced shrinking populations for decades. As small rural towns do not have access to data scientists or even the ability to easily leverage data collected locally to support decisions, our research team will provide communities with data about services in small town Iowa in order to assist with developing strategies to improve quality of life for their residents amid shrinking populations (**?**). We hope to allow towns to discover their own data and compare to other similar towns, centering decision-making on data in the context of small-town Iowa life. In the process, we will assess our visualizations to determine which strategies for user interface and interactive graphics design are most useful to empower town leaders to make discoveries in publicly available data assembled with a focus on items that impact rural quality of life.

2 Data Description

Data collected from data.iowa.gov were used to create the SCC dashboard. Most of these datasets are collected on a town/city or county level, requiring us to carefully join data accounting for differences in spatial resolution. data.iowa.gov contains unique information about residents, including local liquor sales, school building locations, town budgets and expenditures, hospital beds, Medicaid reimbursements, and other details that may provide information about local quality of life. Using this data, we created a dashboard which allows communities to explore these data and compare and contrast their local community to other communities of similar size and location. In addition to manual comparisons created by the user, we will use statistical clustering methods to identify groups of towns which employ similar strategies to maintain resident quality of life.

One of the interesting features of this assembled dataset is that missing data can be missing for multiple reasons: not all state data is complete, but data about certain services may also be missing because towns do not offer that service. Thus, in addition to the usual challenges of working with real-world data that is "messy" in a variety of ways, we also have to contend with missing data that is missing due to the size of the community or the lack of services. This makes both visualization and statistical analysis more complicated.

3 Guiding Design Principles

An additional challenge is that research on dashboard creation and interactive visualization tends to be very task-specific and not generalizable. That is, it is relatively easy to create a dashboard that works for a particular task, but it is hard to generalize from that process what will work for the next dashboard. With this in mind, we have clearly documented our intentions at each stage of the design and evaluation process, with the goal of gathering some useful information about general dashboard design from the process of creating this specific dashboard. Thus, our initial set of dashboard design principles is as follows:

- The town leaders are the focus audience; thus, the town itself should be the central focus of the app.

- Facilitate comparisons with other towns in order to allow the user to explore other potential solutions to offering services that enhance resident quality of life.

- Present the user with peer comparisons in order to widen the scope of exploration beyond the initial set of obvious peers in the local region.

- Allow for more detailed data and feature requests to improve the dashboard design over time.

4 Current Progress and Future Work

In our research, it is important to use visual orientation of dashboards along with the framework of Guided Discovery Learning (GDL), to make sure that the town analyst has the ability to understand the visual orientation easily as well as guided them to new discovery in the dashboard. Stephen Few noted in Information Dashboard Design: The effective visual communication of data the following:

> The visual orientation of dashboards is important due to the speed of perception that's usually required to monitor information. The faster you must assess what's going on, the more you should rely on graphical means to display the information. (**?**)

Few outlined common mistakes that are made in Information Dashboard Design, we will focus on making sure that the following three common mistakes are avoided to make sure that the town analyst have the best path to success:

- Choosing inappropriate display media

- Ineffectively highlighting what's important

- Misusing or overusing color

This list will be the guiding list that can push our work in the direction that is far away from the work done in the workplace in the past. With a great deal of data, many organizations fall into the trap of over ineffectively highlighting what's important along with choosing inappropriate display media and misusing color. This is mostly done, in order to make sure that all of the information is given to the user. This could be a major hurdle when we are collecting a great deal of open-source data for the small towns of Iowa.

A bit of background on Guided Discovery, which is defined as a environment where "students" are actively discovering knowledge. DeDonno stated:

> In our daily lives, we may use a guided discovery process where learning is aided by hints, direction, feedback and other helpful information. (**?**) (**?**)

Our current work will incorporate the ideas of GDL, in which dashboard design will be used to display a panel with vital statistics making sure that the town's analyst small town is the center of focus. The vital statistics will show and highlight the town's financial health, shortest distances from city center to critical resources, such as the fire department, hospital, post offices and public schools. The dashboard will include Quality of Life (QoL) Metrics and the financial health of the small town. Based on the number of variables that are used in making sure that the dashboard is extensive and useful, we will use parallel coordinate plots to show comparisons between the focus small town with the five similar towns.

In the work of the data visualization portion of our research work, a parallel (coordinates) plot allows to compare the feature of several individual observations (series) on a set of numeric variables. Each vertical bar represents a variable and often has its own scale. (The units can even be different). Values are then plotted as series of lines connected across each axis. A parallel coordinate plot is used traditionally in air traffic control to help with collision avoidance algorithms, but has been adapted in the statistical data visualization space with an emphasis on three important considerations: the order, the rotation, and the scaling of the axes.

Using parallel (coordinates) plot will allow

To determine the five similar towns, we will implement unsupervised statistical methods, such as K-means, Principal Component Analysis (PCA), and Hierarchical Clustering, based on distances and services available. We will continue to use clustering methods to identify the towns that will be best compare when the analyst interacts with our dashboard. Using GDL, our dashboard will make sure that the focus small town is compared to other towns to find ways to improve their town using data-driven decisions. Our research partners found out that the small towns do not want to be told what to do but would like to make decisions that are of those who have been apart of the community for generations. Our dashboard will hopefully make sure that the town's analyst are guided in the direction that will help the town's population shrink smart.

In Figure 1, our dashboard will allow the user to select a town name, which will populate the information in the maps related to necessary services, which include directions and distance to fire department, schools, post offices and hospitals. A table will be populated that will include information related to five other towns that have similar attributes, but the aspect of QoL will be included in the differences that help the town make a decision for the town.
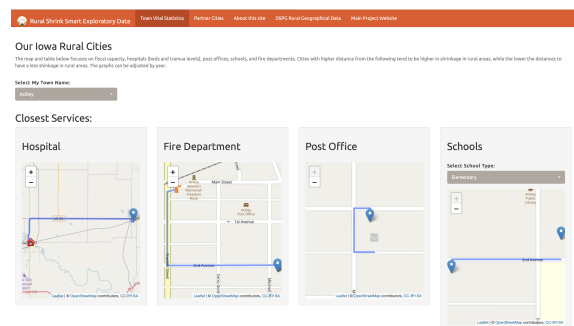

Figure 1: Rural Smart EDA Dashboard Design

4.1 Future Work

We have five towns that have agreed to partner with the project to help with identify the best practices that are useful in the Iowa small and rural towns. As a result to our dashboard development, we will develop a continual feedback loop from the town analyst in these five towns that will help provide our team with useful notes that will be adaptive to all of the small towns in Iowa. Our team plans to capture useful metrics on the use of the product allowing the quantitative and qualitative data collection to allow for a better user experience. This can promote challenges in the event of towns feeling "studied" by people in cities outside of their own.

5 Acknowledgements

References

L. Bartram M. Tory A. Sarikaya, M. Correll and D. Fisher. 2019. What do we talk about when we talk about dashboards? *IEEE*, 25:682–692.

Michael A. DeDonno. 2016. The influence of iq on pure discovery and guided discovery learning of a complex real-world task. *Learning and Individual Differences*, 49:11–16.

Stephen Few. 2006. *Information dashboard design: The effective visual communication of data*. O'Reilly Media, Inc., Newton, MA.

Richard E. Mayer. 2004. Should there be a three-strikes rule against pure discovery learning? *American Psychologist*, 59:14–19.

Rural shrink smart. `https://ruralshrinksmart. org/`. Accessed: 2021-12-01.