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

In our research, it was necessary to use the visual orientation of dashboards and the framework of Guided Discovery Learning (GDL) to make sure that the town analyst can understand the visual orientation easily and guide them to a discovery about possible changes in the dashboard. Stephen Few noted in *Information Dashboard Design: The effective visual communication of data* stated that visual orientation of dashboards is essential due to the speed at which perception (**?**).

A bit of background on Guided Discovery, which is defined as an environment where "students" are actively discovering knowledge, DeDonno states that we use GDL in our daily lives to make decisions by hints, feedback, and helpful information (**?**). Few outlined common mistakes that are made in *Information Dashboard Design*, we have focused on making sure that the following three common mistakes on workplace dashboard development are avoided to best assist the town analyst to the best path of success with our solutions to those common mistakes:

- **Choosing inappropriate display media** - Our work has minimized the number of tables and charts that could come from huge open-source data to incorporate highlighted key metrics and visualize data of five towns in parallel coordinate plots.

- **Ineffectively highlighting what's important** - We wanted the town to be the center of focus for the user; as such, we added their information in the focal point in the center of the page. The comparison towns are only in the parallel coordinate plots to remind users that their town is the most important item in the dashboard.

- **Misusing or overusing color** - Our dashboard uses color to highlight the town in the parallel coordinate plot along with the leaflet maps that will be used in the theme of the Rural Shrink Smart Project

The outlined list has been used as the guiding light to avoid pitfalls in developing dashboards in the workplace. With the amount of data, our dashboard could fall into the trap of over ineffectively highlighting what's important and choosing inappropriate display media and misusing color to ensure that the analyst's attention and all of the information is given to the user. This has been a significant hurdle when displaying the amount of data collected from open-source data for the small towns of Iowa.

Our current work incorporates the ideas of GDL, in which dashboard design has been used to display a panel with vital statistics, ensuring that the town's analyst small town is the center of focus. The vital statistics highlight the town's financial health and the shortest distances from the city center to critical resources, such as the fire department, hospital, post offices, and public schools. The dashboard includes Quality of Life (QoL) Metrics and the financial health of the small town. Based on the number of variables used to ensure that the dashboard is extensive and valuable, we implemented parallel coordinate plots to show comparisons between the focus small town with the five similar towns. A parallel (coordinates) plot allowed our town analyst to compare the essential features of several individual towns on a set of numeric and categorical variables collected. Historically, parallel coordinate plots have been used in air traffic control to help with collision avoidance algorithms but have been adapted in the statistical data visualization space, emphasizing three critical considerations: the order, the rotation, and the scaling of the axes. Our coordinate plots will display variables categorically and numerically by scaling the data.

We have implemented unsupervised statistical methods, such as K-means, Principal Component Analysis (PCA), and Hierarchical Clustering, to determine the five similar towns, based on distances and services available. We continue to use clustering methods to identify the towns that will be best compared when the analyst interacts with our dashboard. Using GDL, our dashboard makes 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 have worked with five towns directly, who have agreed to partner with the project to help identify the best practices that are useful in the small and rural Iowa towns. As a result, we found out that the small towns do not want to be told what to do but would like to make decisions for those who have been a part of the community for generations. Our dashboard will hopefully make sure that the town's analysts are guided in the direction that will help the town's population shrink smart.

Our dashboard development has become more focused on the GDL such that a continual feedback loop from the town analyst in these five towns has helped provide our team with valuable notes that will be adaptive to all of the small towns in Iowa. Our team has captured valuable metrics on using the product allowing the quantitative and qualitative data collection to better user experience. This can promote challenges in towns feeling "studied" by people in cities outside of their own.

In Figure 1, our dashboard allows the user to select a town name, which will populate the information in the maps related to necessary services, including directions and distance to the fire department, schools, post offices, and hospitals. A table populates in the vital statistics sections that have information about the town's QoL Metrics and financial metrics, followed by a parallel coordinate plot that allows the town to see five towns that are similarly based on the most common variables in the towns, such as the distance of fire department.

5 Acknowledgements

We would like to thank our sponsors and grant provider NSF, LSAMP Inspire Program and Iowa League of Cities. We would like to thank our partners and team members at Iowa State University and University of Nebraska - Lincoln. We would like to thank the State of Iowa for providing Data along with our partners in the Rural Iowa communities.

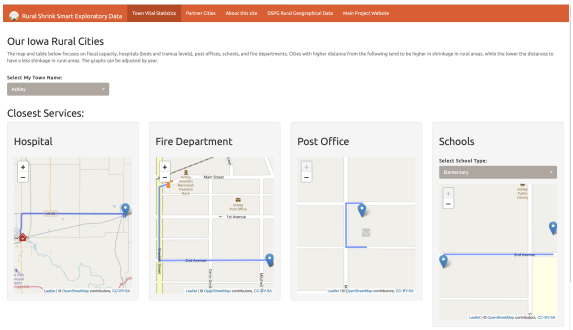

Figure 1: Rural Smart EDA Dashboard Design

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.

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