2023 Data Science Capstone Team

17 May 2023

Final Report

The goal of the capstone project was to work with Southwestern Wisconsin Regional Planning Commission (SWWRP) to create a deliverable report to investigate the causes of failure and trends within Grant, Green, Iowa, Lafeyette, and Richland County municipalities. Our main task was to complete exploratory data analysis to find leading causes and trends for failure, as well as possibly implement machine learning methods to predict future cases. Through multiple meetings with the data sponsors, Troy Maggied and Misty Molzof, we were able to complete a majority of the projected deliverables, as well as many additional ones. The following is a list, in chronological order, of items that we were able to complete as of the time of this report:

* Collection of data from Wisconsin DoR, Dr. Black, and Troy
* Initial Merging of datasets into a consolidated file
* Cleaning of data for inconsistencies
* Exploratory data analysis on multiple factors within the dataset
* Removal of non-important variables as requested by SWWRP
* Defining of initial metric of failure: Revenue – Expenditure
* Refining of metric to encompass the presence of a fund
* Further refinement of second metric to include an actual estimated fund – vulnerability
* Creation of breakout graphs for deeper analysis of counties
* Mockups of GIS maps incorporating the data for spending and vulnerability
* Poster of our findings – used to present to the public
* Calculations made to find percentile of municipality when compared to SW Wisconsin and to the rest of the state
* Visualization of estimate fund over year to see either increase or decrease in spending trends
* Attempted machine learning algorithm
* Final deliverable report which can be used to analyze any municipality in SW Wisconsin
* Final presentation of our findings

Through constant feedback from Troy and Misty, we were able to complete all the items listed above to their satisfaction. Most of the time spent with this project was exploratory data analysis, and refinement of metrics and findings so that they could be presented and eventually implemented into a model. Although our models were unsuccessful given the current data we have, it is possible that with additional data we may be able to predict future failures and vulnerabilities. Additional datasets that may be used for further investigation include actual fund amounts used by municipalities, more up to date population measurements, and additional data on spending by municipalities. There is also room for improvement in terms of the presentation of our project. More definition of terms or project goals and outcomes could lead to a clearer understanding of the project.