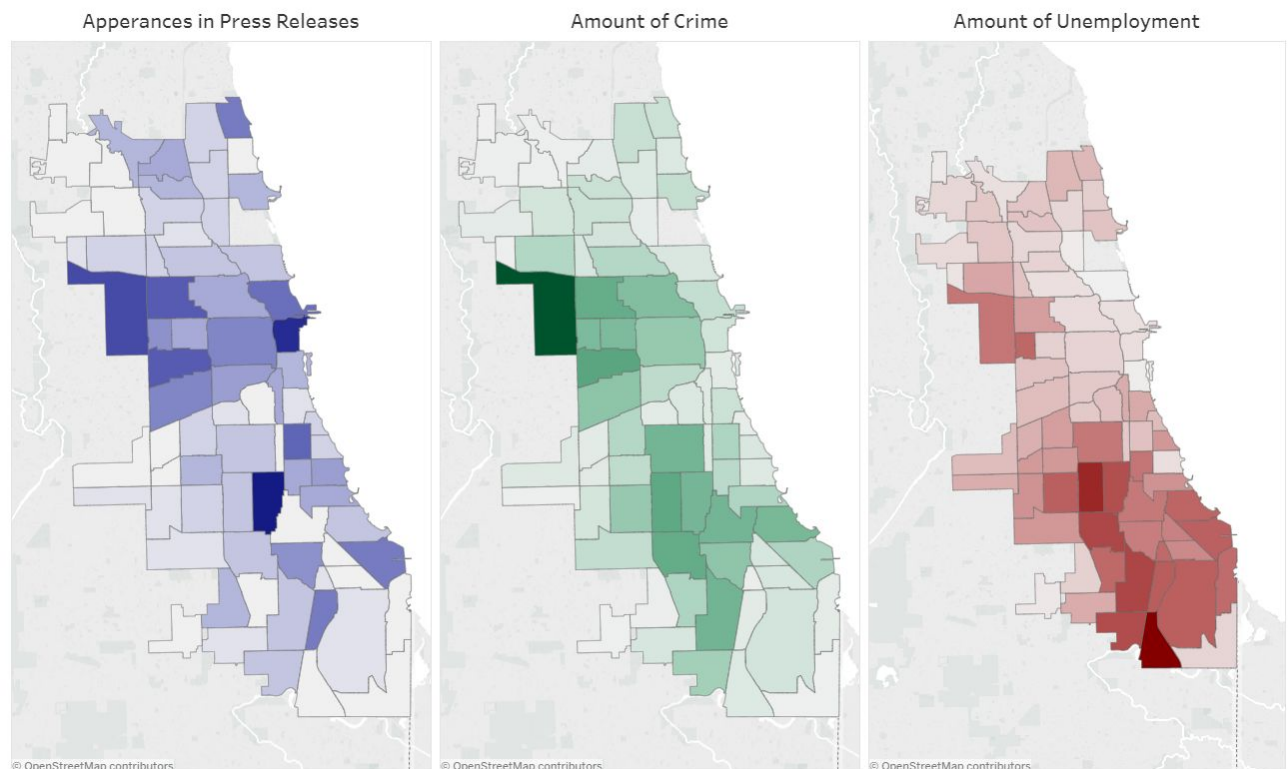


Aileen Nolan
Assessment Report

When looking at the results of this project and taking into consideration how small of a sample my data is, it does appear that there is a relationship between appearances of Chicago Community Area names in press releases published by the mayor's office and community health. The relationship is that Community Areas that score low on community health indicators are generally more talked about in press releases. This was the result I was hoping for and predicted since I considered press releases to be a record of 'what's being done in a city' and how resources are being allocated. In my mind it is a positive thing that the mayor would be talking more about neighborhoods that require more attention and resources.

Below is an image of those results as visualized by Tableau:



What's apparent is that although this relationship can be seen somewhat on the west-side of Chicago, especially in neighborhoods like Humbolt Park and Englewood, the amount of crime has a stronger relationship to appearances in press releases than does unemployment rate.

Considering my project on a whole, some of the drawbacks of my study are that most of my analysis can only be done visually. If I were to take this project one step further, I imagine that some statistical analysis could be done to figure out whether or not there is any strong relationship. Additionally, I would love to expand this project by increasing the sample size of my press releases. My ideal situation would be to repeat this entire project for all of the years that

Mayor Rahm Emanuel has been in office. I wonder how the results might change when you take into consideration the entire course of a mayor's tenure.

Without seeing those types of results, it is hard to say whether or not this project is worth pursuing. I do believe that it was a very good beginner's project because I was not too ambitious in the types of analysis I wanted to do. My main project was trying to figure out the best way to count the appearances of Community Areas within the text. To practice these types of problem-solving skills alone, I believe the project was worth it.

In terms of statistical and data analysis, I still think there is more to be explored. One technique I tried but failed at was using Python packages like Spacy which would recognize text snippets as entities. Getting this type of package to run would absolutely benefit my project. The amount of press releases that were analyzed as having Community Area names was relatively small compared to how many press releases I actually collected. A package like Spacy could help me close the gap and code press releases that don't explicitly call out a Community Area by name but are about some organization, community area, or location that is within these Community Area boundaries. Another beneficial thing of pursuing this would be getting a better grasp of whether the press release data is a fruitful source for actually understanding something about the city. Can we really answer questions from looking at the curated words of the Mayor's Office. How much can we learn and what arguments can we really make from looking at this data? Not too much yet, not until most of the press releases are coded for the content they actually contain. This is believe would definitely be worth exploring further.