hmis: A python tool to visualize and analyze HMIS data

Sara Mahar[1] and Matthew Bellis[1]

1 Siena College

Summary

Many organizations that work to combat homelessness receive funds from the US Department of Housing and Urban Development (HUD). These organizations might be overnight shelters or transitional housing or somewhere in between the Continuum of Care (CoC) provided by these groups. Since 2004, HUD has mandated that groups that receive these funds collect data on the homeless individuals that make use of these services. As such, there is a wealth of data that has been collected all over the country from a variety of organizations. Organizations have some freedom in how they collect and store these data, often making use of 3rd-party software solutions, but the data format is the same everywhere.

This variety of data storage tools means that is is difficult for a data scientist at any of these organizations to dig into this data using standard, open-source computing tools like python or R. These groups can download the data in a standardized "HMIS data dump", which results in 12 separate .csv files, but this still does not make any initial analysis any easier, a priori. These files have information about individuals's name (hashed as a personal ID number), date of birth, prior living, disabilities, jail time, etc.

This module contains a suite of python functions to allow for analysis and visualization of the data collected by the various partners across the CoC. Visualization includes time-series plots, and mapping of the locations of the programs individuals have entered. Analysis can be done with these visualizations and with the ability to withdraw individuals who share a common character. For example, the analyst can withdraw all of the individuals who have visited more than 25 programs and then visualize them.

We have developed these tools to work with the standard HMIS data dump in the RHY (Runaway and Homeless Youth) data format, that produces 12 .csv files in which personal identifying information is de-identified through a hashing algorithm. Because of this standardization, any other tools that leverage this software package can be used by similar networks across the country.

The definitions of the information in the HMIS data can be found on HUD's website (HUD 2017 ; HUD 2016).

References

HUD, US Department of Housing and Urban Development. 2016. "HMIS Data Standards Data Manual." US Government. https://www.hudexchange.info/resources/documents/HMIS-Data-Standards-Manual.pdf.

———. 2017. "HMIS Data Standards Data Dictionary." US Government. https://www.hudexchange.info/resource/3824/hmis-data-dictionary/.