The Journal of Open Source Software

findSVI: an R package to calculate the Social Vulnerability Index at multiple geographical levels

Heli Xu[1¶], **Ran Li**[1], and **Usama Bilal**[1,2]

1 Urban Health Collaborative, Drexel Dornsife School of Public Health, Philadelphia, PA, United States of America **2** Department of Epidemiology and Biostatistics, Drexel Dornsife School of Public Health, Philadelphia, PA, United States of America ¶ Corresponding author

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Øystein Sørensen

Reviewers:

- @epiben

Submitted: 04 March 2024
Published: unpublished

Summary

The Social Vulnerability Index (SVI) was created by the Centers for Disease Control and Prevention and Agency for Toxic Substances and Disease Registry (CDC/ATSDR) as a tool to assess the resilience of communities in preparation for public health crisis (Flanagan et al., 2011). In the event of public health emergencies, communities with higher SVI are considered more vulnerable, i.e., more likely to suffer serious consequences such as mortality/physical injuries, displacement and economic loss. The Geospatial Research, Analysis, and Services Program (GRASP) maintains the SVI database and updates the SVI every two years (Disease Control et al., 2021).

The SVI uses 16 variables from the American Community Survey (ACS) (Flanagan et al., 2011).These 16 variables are grouped into 4 themes/domains: Socioeconomic Status, Household Characteristics, Racial & Ethnic Minority Status and Housing Type & Transportation. For each variable, percentile ranks are calculated and aggregated (summed) for each theme and all themes for a geographic level of interest, which are used to generate another set of percentile ranks for a geographic unit as the theme-specific and overall SVIs. The CDC/ATSDR provides the SVI at the census tract and county levels, with percentiles calculated with respect to each state and the entire United States of America.

The main aim of the findSVI R package is to extend the application of CDC/ATSDR SVI methodology to more geographic levels and to utilize more up-to-date ACS data. Leveraging the tidycensus R package (Walker & Herman, 2024), findSVI allows data retrieval from American Community Survey for any year from 2012 to 2021, with more options for geographic levels (e.g., zip code tabulation areas or ZCTAs, places) and regions of reference for percentile calculations (e.g., county-specific, multiple states). In addition, findSVI provides an efficient SVI analysis workflow for requests involving multiple year-state pairs that need to be ranked separately. Apart from returning the result as an SVI data frame, findSVI also supports output as an SVI table with simple feature geometry for spatial analysis in R and other Geographic Information Systems tools.

Statement of Need

The SVI is a widely used indicator to measure the differential impact of public health crisis on communities with different socioeconomic and demographic characteristics. In addition to facilitating effective planning of social services and public assistance for disasters (Flanagan et al., 2018; Schmidt et al., 2021), the SVI has also provided a valuable metric for characterizing the relationships between social vulnerability and various health outcomes, such as surgical outcomes (Hyer et al., 2021; Paro et al., 2021), environmental exposure-related illness (Do et al., 2023; Lehnert et al., 2020), cardiovascular diseases (Bevan et al., 2022; Ibrahim et al., 2023; Khan et al., 2021), as well as COVID-19 in recent years (Hughes et al., 2021; Karaye & Horney, 2020; Khazanchi et al., 2020; Tsai et al., 2022). In particular, CDC/ATSDR SVI has been used to examine spatial and racial/ethnic inequities in COVID-19 resources and the disproportionate impact of COVID-19 on vulnerable communities (Bilal et al., 2021; Bilal, Jemmott, et al., 2022; Bilal, Mullachery, et al., 2022; Gaynor & Wilson, 2020; Islam et al., 2021; Mullachery et al., 2022).

However, many of the applications above had to recalculate the census-tract level SVI to match their units of analysis (Bilal et al., 2021; Bilal, Jemmott, et al., 2022; Bilal, Mullachery, et al., 2022), including ZCTAs, a unit of analysis for which public health data tends to be available quicker in public health emergencies (Bilal et al., 2021). With findSVI, the workflow involved in obtaining SVI can be streamlined and standardized for a specific research question. Provided with year(s), region(s) of reference and a geographic level of interest, findSVI retrieves the required variables from the US Census Bureau data and calculates SVI for communities in the specified area. Furthermore, with its features to retrieve ACS data at more flexible geographic levels and for more recent years, findSVI offers researchers more options for using SVI to accurately reflect the region/year of focus for both association studies and spatial analyses.

Functionality

Basic usage of retrieving Census data and computing SVI for a region at a geographic level can be found in the vignette Introduction to findSVI. In addition, Validation of SVI Results contains a reference table of the SVI variables and ACS variables across available years and a comparison between SVI database and findSVI calculations. Examples involving spatial analyses with SVI are included in Mapping SVI for Spatial Analysis and Geographic Context of SVI. An example of using findSVI to study social vulnerabilities and health outcomes is also included in SVI and health outcome, where the association between SVI and COVID-19 hospitalizations in Philadelphia is explored with maps and a scatter plot.

FindSVI makes use of the tidycensus package (Walker & Herman, 2024) for Census data retrieval, dplyr (Wickham et al., 2023), purrr (Wickham & Henry, 2023), stringr (Wickham, 2023), tidyr (Wickham et al., 2024), tidyselect (Henry & Wickham, 2024b), rlang (Henry & Wickham, 2024a) packages for data manipulation and cli (Csárdi, 2023) package for user-friendly error messages.

References

Bevan, G., Pandey, A., Griggs, S., Dalton, J. E., Zidar, D., Patel, S., Khan, S. U., Nasir, K., Rajagopalan, S., & Al-Kindi, S. (2022). Neighborhood-level social vulnerability and prevalence of cardiovascular risk factors and coronary heart disease. *Current Problems in Cardiology*, 101182. https://doi.org/10.1016/j.cpcardiol.2022.101182

Bilal, U., Jemmott, J. B., Schnake-Mahl, A., Murphy, K., & Momplaisir, F. (2022). Racial/ethnic and neighbourhood social vulnerability disparities in COVID-19 testing positivity, hospitalization, and in-hospital mortality in a large hospital system in pennsylvania: A prospective study of electronic health records. *The Lancet Regional Health–Americas*, *10*. https://doi.org/10.1016/j.lana.2022.100220

Bilal, U., Mullachery, P. H., Schnake-Mahl, A., Rollins, H., McCulley, E., Kolker, J., Barber, S., & Diez Roux, A. V. (2022). Heterogeneity in spatial inequities in COVID-19 vaccination across 16 large US cities. *American Journal of Epidemiology*, *191*(9), 1546–1556. https://doi.org/10.1093/aje/kwac076

Bilal, U., Tabb, L. P., Barber, S., & Diez Roux, A. V. (2021). Spatial inequities in COVID-19 testing, positivity, confirmed cases, and mortality in 3 US cities: An ecological study. *Annals of Internal Medicine*, *174*(7), 936–944. https://doi.org/10.7326/M20-3936

Csárdi, G. (2023). *Cli: Helpers for developing command line interfaces*. https://cli.r-lib.org

Disease Control, C. for, Toxic Substances, P. A. for, Disease Registry/ Geospatial Research, Analysis, & Program, S. (2021). *CDC/ATSDR SVI data and documentation download*. https://www.atsdr.cdc.gov/placeandhealth/svi/data_documentation_download.html

Do, V., McBrien, H., Flores, N. M., Northrop, A. J., Schlegelmilch, J., Kiang, M. V., & Casey, J. A. (2023). Spatiotemporal distribution of power outages with climate events and social vulnerability in the USA. *Nature Communications*, *14*(1), 2470. https://doi.org/https://doi.org/10.1038/s41467-023-38084-6

Flanagan, B. E., Gregory, E. W., Hallisey, E. J., Heitgerd, J. L., & Lewis, B. (2011). A social vulnerability index for disaster management. *Journal of Homeland Security and Emergency Management*, *8*(1), 0000102202154773551792. https://doi.org/https://doi.org/10.2202/1547-7355.1792

Flanagan, B. E., Hallisey, E. J., Adams, E., & Lavery, A. (2018). Measuring community vulnerability to natural and anthropogenic hazards: The centers for disease control and prevention's social vulnerability index. *Journal of Environmental Health*, *80*(10), 34. https://pubmed.ncbi.nlm.nih.gov/32327766/

Gaynor, T. S., & Wilson, M. E. (2020). Social vulnerability and equity: The disproportionate impact of COVID-19. *Public Administration Review*, *80*(5), 832–838. https://doi.org/10.1111/puar.13264

Henry, L., & Wickham, H. (2024a). *Rlang: Functions for base types and core r and 'tidyverse' features*. https://rlang.r-lib.org

Henry, L., & Wickham, H. (2024b). *Tidyselect: Select from a set of strings*. https://tidyselect.r-lib.org

Hughes, M. M., Wang, A., Grossman, M. K., Pun, E., Whiteman, A., Deng, L., Hallisey, E., Sharpe, J. D., Ussery, E. N., Stokley, S., & others. (2021). County-level COVID-19 vaccination coverage and social vulnerability—united states, december 14, 2020–march 1, 2021. *Morbidity and Mortality Weekly Report*, *70*(12), 431. https://doi.org/http://dx.doi.org/10.15585/mmwr.mm7012e1

Hyer, J. M., Tsilimigras, D. I., Diaz, A., Mirdad, R. S., Azap, R. A., Cloyd, J., Dillhoff, M., Ejaz, A., Tsung, A., & Pawlik, T. M. (2021). High social vulnerability and "textbook outcomes" after cancer operation. *Journal of the American College of Surgeons*, *232*(4), 351–359. https://doi.org/10.1016/j.jamcollsurg.2020.11.024

Ibrahim, R., Salih, M., Gomez Tirambulo, C. V., Takamatsu, C., Lee, J. Z., Fortuin, D., & Lee, K. S. (2023). Impact of social vulnerability and demographics on ischemic heart disease mortality in the united states. *JACC: Advances*, *2*(7), 100577. https://doi.org/https://doi.org/10.1016/j.jacadv.2023.100577

Islam, S. J., Nayak, A., Hu, Y., Mehta, A., Dieppa, K., Almuwaqqat, Z., Ko, Y.-A., Patel, S. A., Goyal, A., Sullivan, S., & others. (2021). Temporal trends in the association of social vulnerability and race/ethnicity with county-level COVID-19 incidence and outcomes in the USA: An ecological analysis. *BMJ Open*, *11*(7), e048086. https://doi.org/10.1136/bmjopen-2020-048086

Karaye, I. M., & Horney, J. A. (2020). The impact of social vulnerability on COVID-19 in the US: An analysis of spatially varying relationships. *American Journal of Preventive Medicine*, *59*(3), 317–325. https://doi.org/10.1016/j.amepre.2020.06.006

Khan, S. U., Javed, Z., Lone, A. N., Dani, S. S., Amin, Z., Al-Kindi, S. G., Virani, S. S., Sharma, G., Blankstein, R., Blaha, M. J., & others. (2021). Social vulnerability and premature cardiovascular mortality among US counties, 2014 to 2018. *Circulation*, *144*(16), 1272–1279. https://doi.org/https://doi.org/10.1161/CIRCULATIONAHA.121.054516

Khazanchi, R., Beiter, E. R., Gondi, S., Beckman, A. L., Bilinski, A., & Ganguli, I. (2020). County-level association of social vulnerability with COVID-19 cases and deaths in the USA. *Journal of General Internal Medicine*, *35*, 2784–2787. https://doi.org/10.1007/s11606-020-05882-3

Lehnert, E. A., Wilt, G., Flanagan, B., & Hallisey, E. (2020). Spatial exploration of the CDC's social vulnerability index and heat-related health outcomes in georgia. *International Journal of Disaster Risk Reduction*, *46*, 101517. https://doi.org/10.1016/j.ijdrr.2020.101517

Mullachery, P. H., Li, R., Melly, S., Kolker, J., Barber, S., Roux, A. V. D., & Bilal, U. (2022). Inequities in spatial accessibility to COVID-19 testing in 30 large US cities. *Social Science & Medicine*, *310*, 115307. https://doi.org/10.1016/j.socscimed.2022.115307

Paro, A., Hyer, J. M., Diaz, A., Tsilimigras, D. I., & Pawlik, T. M. (2021). Profiles in social vulnerability: The association of social determinants of health with postoperative surgical outcomes. *Surgery*, *170*(6), 1777–1784. https://doi.org/10.1016/j.surg.2021.06.001

Schmidt, H., Weintraub, R., Williams, M. A., Miller, K., Buttenheim, A., Sadecki, E., Wu, H., Doiphode, A., Nagpal, N., Gostin, L. O., & others. (2021). Equitable allocation of COVID-19 vaccines in the united states. *Nature Medicine*, *27*(7), 1298–1307. https://doi.org/https://doi.org/10.1038/s41591-021-01379-6

Tsai, T. C., Jacobson, B. H., Orav, E. J., & Jha, A. K. (2022). Association of community-level social vulnerability with US acute care hospital intensive care unit capacity during COVID-19. *Healthcare*, *10*, 100611. https://doi.org/https://doi.org/10.1016/j.hjdsi.2021.100611

Walker, K., & Herman, M. (2024). *Tidycensus: Load US census boundary and attribute data as 'tidyverse' and 'sf'-ready data frames*. https://walker-data.com/tidycensus/

Wickham, H. (2023). *Stringr: Simple, consistent wrappers for common string operations*. https://stringr.tidyverse.org

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *Dplyr: A grammar of data manipulation*. https://dplyr.tidyverse.org

Wickham, H., & Henry, L. (2023). *Purrr: Functional programming tools*. https://purrr.tidyverse.org/

Wickham, H., Vaughan, D., & Girlich, M. (2024). *Tidyr: Tidy messy data*. https://tidyr.tidyverse.org