

Package ‘ndi’

August 13, 2022

Title Neighborhood Deprivation Indices

Version 0.1.0

Date 2022-08-10

Maintainer Ian D. Buller <ian.buller@nih.gov>

Description Compute various neighborhood deprivation indices (NDI), including:
(1) based on Messer et al. (2006) <[doi:10.1007/s11524-006-9094-x](https://doi.org/10.1007/s11524-006-9094-x)>
and (2) based on Andrews et al. (2020) <[doi:10.1080/17445647.2020.1750066](https://doi.org/10.1080/17445647.2020.1750066)>
and Slotman et al. (2022) <[doi:10.1016/j.dib.2022.108002](https://doi.org/10.1016/j.dib.2022.108002)>
who uses variables chosen by Roux and Mair (2010)
<[doi:10.1111/j.1749-6632.2009.05333.x](https://doi.org/10.1111/j.1749-6632.2009.05333.x)>. Both are a decomposition
of multiple demographic characteristics from the U.S. Census Bureau
American Community Survey 5-year estimates.

License Apache License (>= 2.0)

Encoding UTF-8

RoxygenNote 7.2.1

Depends R (>= 3.5.0)

Imports dplyr, MASS, psych, stats, stringr, tidycensus, tidyr

Suggests ggplot2, testthat, tigris, R.rsp, sf, spelling

VignetteBuilder R.rsp

Language en-US

URL <https://github.com/idblr/ndi>

BugReports <https://github.com/idblr/ndi/issues>

NeedsCompilation no

Author Ian D. Buller [aut, cre, cph] (<<https://orcid.org/0000-0001-9477-8582>>),
NCI [cph, fnd]

Repository CRAN

Date/Publication 2022-08-13 12:00:02 UTC

R topics documented:

ndi-package	2
messer	3
powell_wiley	5

Index	8
--------------	----------

ndi-package	<i>The ndi Package: Neighborhood Deprivation Indices</i>
-------------	--

Description

Computes various geospatial neighborhood deprivation indices in the United States

Details

The 'ndi' package computes various neighborhood deprivation indices (NDI), including: (1) based on Messer et al. (2006) [doi:10.1007/s115240069094x](https://doi.org/10.1007/s115240069094x) and (2) based on Andrews et al. (2020) [doi:10.1080/17445647.2020.1750066](https://doi.org/10.1080/17445647.2020.1750066) and Slotman et al. (2022) [doi:10.1016/j.dib.2022.108002](https://doi.org/10.1016/j.dib.2022.108002) who uses variables chosen by Roux and Mair (2010) [doi:10.1111/j.17496632.2009.05333.x](https://doi.org/10.1111/j.17496632.2009.05333.x). Both are a decomposition of multiple demographic characteristics from the U.S. Census Bureau American Community Survey 5-year estimates.

Key content of the 'ndi' package include:

Neighborhood Deprivation Indices

[messer](#) Computes NDI values based on Messer et al. (2006) [doi:10.1007/s115240069094x](https://doi.org/10.1007/s115240069094x).

[powell_wiley](#) Computes NDI values based on Andrews et al. (2020) [doi:10.1080/17445647.2020.1750066](https://doi.org/10.1080/17445647.2020.1750066) and Slotman et al. (2022) [doi:10.1016/j.dib.2022.108002](https://doi.org/10.1016/j.dib.2022.108002) who uses variables chosen by Roux and Mair (2010) [doi:10.1111/j.17496632.2009.05333.x](https://doi.org/10.1111/j.17496632.2009.05333.x).

Dependencies

The 'ndi' package relies heavily upon [tidycensus](#) and [psych](#) for computing the neighborhood deprivation indices. The [messer](#) function builds upon code developed by Hruska et al. (2022) [doi:10.17605/OSF.IO/M2SAV](https://doi.org/10.17605/OSF.IO/M2SAV) by fictionalizing, adding percent households earning <\$30,000 per year to the NDI computation, and providing the option for computing the ACS-5 2006-2010 NDI values. There was no code companion to compute NDI included in Andrews et al. (2020) [doi:10.1080/17445647.2020.1750066](https://doi.org/10.1080/17445647.2020.1750066) or Slotman et al. (2022) [doi:10.1016/j.dib.2022.108002](https://doi.org/10.1016/j.dib.2022.108002), but the package maintainer worked directly with the authors to replicate their SAS code in R.

Author(s)

Ian D. Buller

Occupational and Environmental Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, Maryland, USA.

Maintainer: I.D.B. <ian.buller@nih.gov>

 messer

Neighborhood Deprivation Index based on Messer et al. (2006)

Description

Compute the Neighborhood Deprivation Index (Messer) values.

Usage

```
messer(geo = "tract", year = 2020, imp = FALSE, quiet = FALSE, ...)
```

Arguments

geo	Character string specifying the geography of the data either census tracts geo = "tract" (the default) or counties geo = "county".
year	Numeric. The year to compute the estimate. The default is 2020 and the years between 2010 and 2020 are currently available.
imp	Logical. If TRUE, will impute missing census characteristics within the internal principal . If FALSE (the default), will not impute.
quiet	Logical. If TRUE, will display messages about potential missing census information and proportion of variance explained by principal component analysis. The default is FALSE.
...	Arguments passed to get_acs to select state, county, and other arguments for census characteristics

Details

This function will compute the Neighborhood Deprivation Index (NDI) of U.S. census tracts or counties for a specified geographical referent (e.g., US-standardized) based on Messer et al. (2006) [doi:10.1007/s115240069094x](https://doi.org/10.1007/s115240069094x).

The function uses the [get_acs](#) function to obtain U.S. Census Bureau 5-year American Community Survey characteristics used for computation involving a principal component analysis with the [principal](#) function. The yearly estimates are available 2010 and after when all census characteristics became available. The eight characteristics are:

- C24030: percent males in management, science, and arts occupation
- B25014: percent of crowded housing
- B17017: percent of households in poverty
- B25115: percent of female headed households with dependents
- B19058: percent of households on public assistance
- B19001: percent households earning <\$30,000 per year
- B06009: percent earning less than a high school education
- B23025: percent unemployed (2011 onward)

- B23001: percent unemployed (2010 only)

Use the internal state and county arguments within the `get_acs` function to specify the referent for standardizing the NDI (Messer) values. For example, if all U.S. states are specified for the state argument, then the output would be a U.S.-standardized index.

The continuous NDI (Messer) values are z-transformed, i.e., "standardized," and the categorical NDI (Messer) values are quartiles of the standardized continuous NDI (Messer) values.

Check if the proportion of variance explained by the first principal component is high (more than 0.5).

Value

An object of class 'list'. This is a named list with the following components:

`ndi` An object of class 'tbl' for the GEOID, name, NDI (standardized), NDI (quartile), and raw census values of specified census tracts.

`pca` An object of class 'principal', returns the output of `principal` used to compute the NDI values.

`missing` An object of class 'tbl' of the count and proportion of missingness for each census variable used to compute the NDI.

See Also

`get_acs` for additional arguments for geographic referent selection (i.e., state and county).

Examples

```
## Not run:
# Wrapped in \dontrun{} because these examples require a Census API key.

# Tract-level metric (2020)
messer(geo = "tract", state = "GA", year = 2020)

# Impute NDI for tracts (2020) with missing census information (median values)
messer(state = "tract", "GA", year = 2020, imp = TRUE)

# County-level metric (2020)
messer(geo = "county", state = "GA", year = 2020)

## End(Not run)
```

powell_wiley	<i>Neighborhood Deprivation Index based on Andrews et al. (2020) and Slotman et al. (2022)</i>
--------------	--

Description

Compute the Neighborhood Deprivation Index (Powell-Wiley) values.

Usage

```
powell_wiley(geo = "tract", year, imp = FALSE, quiet = FALSE, ...)
```

Arguments

geo	Character string specifying the geography of the data either census tracts <code>geo = "tract"</code> (the default) or counties <code>geo = "county"</code> .
year	Numeric. The year to compute the estimate (2010-2020 currently available).
imp	Logical. If TRUE, will impute missing census characteristics within the internal principal using median values of variables. If FALSE (the default), will not impute.
quiet	Logical. If TRUE, will display messages about potential missing census information, standardized Cronbach's alpha, and proportion of variance explained by principal component analysis. The default is FALSE.
...	Arguments passed to get_acs to select state, county, and other arguments for census characteristics

Details

This function will compute the Neighborhood Deprivation Index (NDI) of U.S. census tracts or counties for a specified geographical referent (e.g., US-standardized) based on Andrews et al. (2020) [doi:10.1080/17445647.2020.1750066](https://doi.org/10.1080/17445647.2020.1750066) and Slotman et al. (2022) [doi:10.1016/j.dib.2022.108002](https://doi.org/10.1016/j.dib.2022.108002).

The function uses the [get_acs](#) function to obtain U.S. Census Bureau 5-year American Community Survey characteristics used for computation involving a factor analysis with the [principal](#) function. The yearly estimates are available 2010 and after when all census characteristics became available. The thirteen characteristics chosen by Roux and Mair (2010) [doi:10.1111/j.1749-6632.2009.05333.x](https://doi.org/10.1111/j.1749-6632.2009.05333.x) are:

- MedHHInc (5B19013): median household income (dollars)
- PctRecvIDR (B19054): percent of households receiving dividends, interest, or rental income
- PctPubAsst (B19058): percent of households receiving public assistance
- MedHomeVal (B25077): median home value (dollars)
- PctMgmtBusScArti (C24060): percent in a management, business, science, or arts occupation
- PctFemHeadKids (B11005): percent of households that are female headed with any children under 18 years

- PctOwnerOcc (DP04): percent of housing units that are owner occupied
- PctNoPhone (DP04): percent of households without a telephone
- PctNComPlm (DP04): percent of households without complete plumbing facilities
- PctEducHSPlus (S1501): percent with a high school degree or higher (population 25 years and over)
- PctEducBchPlus (S1501): percent with a college degree or higher (population 25 years and over)
- PctFamBelowPov (S1702): percent of families with incomes below the poverty level
- PctUnempl (S2301): percent unemployed

Use the internal state and county arguments within the `get_acs` function to specify the referent for standardizing the NDI (Powell-Wiley) values. For example, if all U.S. states are specified for the state argument, then the output would be a U.S.-standardized index. Please note: the NDI (Powell-Wiley) values will not exactly match (but will highly correlate with) those found in Andrews et al. (2020) [doi:10.1080/17445647.2020.1750066](https://doi.org/10.1080/17445647.2020.1750066) and Slotman et al. (2022) [doi:10.1016/j.dib.2022.108002](https://doi.org/10.1016/j.dib.2022.108002) because the two studies used a different statistical platform (i.e., SPSS and SAS, respectively) that intrinsically calculate the principal component analysis differently from R.

The categorical NDI (Powell-Wiley) values are population-weighted quintiles of the continuous NDI (Powell-Wiley) values.

Check if the proportion of variance explained by the first principal component is high (more than 0.5).

Value

An object of class 'list'. This is a named list with the following components:

`ndi` An object of class 'tbl' for the GEOID, name, NDI continuous, NDI quintiles, and raw census values of specified census tracts.

`pca` An object of class 'principal', returns the output of `principal` used to compute the NDI values.

`missing` An object of class 'tbl' of the count and proportion of missingness for each census variable used to compute the NDI.

`cronbach` An object of class 'character' or 'numeric' for the results of the Cronbach's alpha calculation. If only one factor is computed, a message is returned. If more than one factor is computed, the Cronbach's alpha is calculated and should check that it is >0.7 for respectable internal consistency between factors.

See Also

`get_acs` for additional arguments for geographic referent selection (i.e., state and county).

Examples

```
## Not run:
# Wrapped in \dontrun{} because these examples require a Census API key.

# Tract-level metric (2020)
powell_wiley(geo = "tract", state = "GA", year = 2020)
```

```
# Impute NDI for tracts (2020) with missing census information (median values)
powell_wiley(state = "tract", "GA", year = 2020, imp = TRUE)

# County-level metric (2020)
powell_wiley(geo = "county", state = "GA", year = 2020)

## End(Not run)
```

Index

*** package**

ndi-package, 2

get_acs, 3–6

messer, 2, 3

ndi (ndi-package), 2

ndi-package, 2

powell_wiley, 2, 5

principal, 3–6

psych, 2

tidycensus, 2