Package: zctaDB (via r-universe)

October 29, 2024

Title Get Geomarker Data Based on ZIP Code

Version 0.1.0

Description Get geomarker data for ZIP Codes in the contiguous United States. Available data includes racial composition; length and density of primary and secondary roads' land cover classifications including greenness, tree canopy, and imperviousness; measures of community deprivation; NARR cell identifier; and data from the EPA environmental justice index including traffic proximity, ozone concentration, PM2.5 concentration, concentration of PM from diesel traffic, and a respiratory hazard index.

License GPL-3

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.2.1

Depends R (>= 2.10)

Imports dplyr

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

Repository https://geomarker-io.r-universe.dev

RemoteUrl https://github.com/geomarker-io/zctaDB

RemoteRef HEAD

RemoteSha 73e36037e9eb7f9b99addb5098644532718a0e66

Contents

Index   **8**

---

add_aadt_data   *Get Traffic Data*

Description

adds Average Annual Daily Traffic Density. AADT data is from 2017 and corresponds to 2010 ZCTA vintages.

Usage

```
add_aadt_data(data)
```

Arguments

data   data.frame or tibble with column called 'zcta' at minimum.

Value

the input data.frame with the following columns appended

moving_roads_density density of roads with moving traffic (meters of road per square meter of area)

stop_go_roads_density density of roads with stop and go traffic (meters of road per square meter of area)

moving_traffic_density moving traffic density (vehicle-meters per square meter of area)

stop_go_traffic_density stop and go traffic density (vehicle-meters per square meter of area)

moving_truck_density moving truck traffic density (truck-meters per square meter of area)

stop_go_truck_density stop and go truck traffic density (truck-meters per square meter of area)

References

https://degauss.org/aadt

Examples

```
my_data <- data.frame(zcta = c('45229', '45056', '47012'))
add_aadt_data(data = my_data)
```

---

add_depindex_data   *Get Community Deprivation Data*

Description

adds community deprivation index and related variables from the American Community Survey (ACS). Data is available for 2000 and 2010 ZCTAs and for years 2015 and 2018.

Usage

```
add_depindex_data(data)
```

Arguments

data   data.frame or tibble with column called 'zcta' at minimum. Optionally, another column called 'year', with possible values 2000 to 2019. If no year is given, defaults to 2018

Value

the input data.frame with the following columns appended

zcta_year the ZCTA boundary year matched to the input year (2000 or 2010)

dep_year the ACS data year matched to the input year (2015 or 2018)

fraction_assisted_income fraction of households receiving public assistance income or food stamps or SNAP in the past 12 months

fraction_high_school_edu fraction of population 25 and older with educational attainment of at least high school graduation (includes GED equivalency)

median_income median household income in the past 12 months in 2015 inflation-adjusted dollars

fraction_no_health_ins fraction of poulation with no health insurance coverage

fraction_poverty fraction of population with income in past 12 months below poverty level

fraction_vacant_housing fraction of houses that are vacant

dep_index a composite measure of the other 6 variables

References

https://geomarker.io/dep_index/

Examples

```
my_data <- data.frame(zcta = c('45229', '45056', '47012'))
add_depindex_data(data = my_data)
```

`add_ejscreen_data` *Get EJ Screen Data*

Description

adds traffic proximity index, ozone concentration, particulate matter (PM) concentration, PM from diesel traffic, and a respiratory hazard index from the EPA EJ Screen database. Data is available for 2000 and 2010 ZCTAs and for years 2015 to 2018.

Usage

```
add_ejscreen_data(data)
```

Arguments

data   data.frame or tibble with column called 'zcta' at minimum. Optionally, another column called 'year', with possible values 2000 to 2019. If no year is given, defaults to 2018.

Value

the input data.frame with the following columns appended

`zcta_year` the ZCTA boundary year matched to the input year (2000 or 2010)

`ej_year` the EJ Screen data year matched to the input year (2015 - 2018)

`traffic_proximity` count of vehicles (AADT, avg. annual daily traffic) at major roads within 500 meters, divided by distance in meters

`ozone_conc` ozone summer seasonal avg. of daily maximum 8-hour concentration in air in parts per billion

`pm_conc` PM2.5 levels in air, μg/m3 annual avg

`diesel_pm` diesel particulate matter level in air, ug/m3

`resp_hazard_ind` air toxics respiratory hazard index (ratio of exposure concentration to health-based reference concentration)

References

https://www.epa.gov/ejscreen/overview-environmental-indicators-ejscreen

Examples

```
my_data <- data.frame(zcta = c('45229', '45056', '47012'))
add_ejscreen_data(data = my_data)
```

`add_landcover_data` *Get Land Cover Data*

Description

adds percentage of land classified as 'green', percentage impervious land, and percentage of land covered by tree canopy from the National Landcover Database. Data is available for 2000 and 2010 ZCTAs and for years 2001, 2006, 2011, and 2016.

Usage

```
add_landcover_data(data)
```

Arguments

data   data.frame or tibble with column called 'zcta' at minimum. Optionally, another column called 'year', with possible values 2000 to 2019. If no year is given, defaults to 2016.

Value

the input data.frame with the following columns appended

`zcta_year` the ZCTA boundary year matched to the input year (2000 or 2010)

`nlcd_year` the National Landcover data year matched to the input year (2001, 2006, 2011, or 2016)

`treecanopy_year` if tree canopy data was not available in the nlcd_year, the year of the tree canopy data, otherwise NA

`pct_green` percentage of pixels within ZCTA classified as green (defined as all NLCD land cover codes except water, ice/snow, developed medium intensity, developed high intensity, and rock/sand/clay)

`pct_impervious` average percent impervious for pixels within ZCTA

`pct_treecanopy` average percent tree canopy for pixels within ZCTA

References

https://www.mrlc.gov/

Examples

```
my_data <- data.frame(zcta = c('45229', '45056', '47012'))
add_landcover_data(data = my_data)
```

---

add_narr_cell_zcta   *Get NARR Cell ID for ZCTA Centroid*

Description

adds NARR cell identifier based on ZCTA centroid for subsequent use with the addNarrData package-age

Usage

```
add_narr_cell_zcta(data)
```

Arguments

data   data.frame or tibble with column called 'zcta' at minimum.

Value

the input data.frame with the following columns appended

narr_cell NARR cell indentifier corresponding to the NARR cell that intersects the ZCTA centroid

References

https://geomarker.io/addNarrData/

Examples

```
my_data <- data.frame(zcta = c('45229', '45056', '47012'))
add_narr_cell_zcta(data = my_data)
```

---

add_race_data   *Get Racial Composition Data*

Description

adds numbers and percentages of white non-Hispanic and black non-Hispanic residents from the 2018 ACS for given ZIP codes, as well as a racial index of concentration at the extremes (ICE). Data is available for 2010 ZCTAs in the contiguous United States.

Usage

```
add_race_data(data)
```

Arguments

data   data.frame or tibble with column called 'zcta' at minimum.

Value

the input data.frame with the following columns appended

`total` total population of ZCTA

wnh number of white non-Hispanic residents

bnh number of black non-Hispanic residents

`racial_ice` racial index of concentration at the extremes (wnh-bnh)/total)

`pct_wnh` percent white wnh/totalx100

`pct_bnh` percent black bnh/totalx100

Examples

```
my_data <- data.frame(zcta = c('45229', '45056', '47012'))
add_race_data(data = my_data)
```

---

add_road_data   *Get Length and Density of Roads*

Description

adds length (m) and density (m/km2) of primary and secondary roads as defined by 2018 TIGER Line files. Data is available for 2000 and 2010 ZCTAs.

Usage

```
add_road_data(data)
```

Arguments

data   data.frame or tibble with column called 'zcta' at minimum. Optionally, another column called 'year', with possible values 2000 to 2019. Years 2000 to 2009 will be matched to 2000 ZCTAs and years 2010 to 2019 will be matched to 2010 ZCTAs. If no year is given, defaults to 2018.

Value

the input data.frame with the following columns appended

`primary_road_length` length (m) of primary roads in ZCTA

`primary_road_density` length(m) of primary roads in ZCTA divided by area of ZCTA (km2)

`secondary_road_length` length (m) of secondary roads in ZCTA

`secondary_road_density` length(m) of secondary roads in ZCTA divided by area of ZCTA (km2)

Examples

```
my_data <- data.frame(zcta = c('45229', '45056', '47012'))
add_road_data(data = my_data)
```

Index