Package 'fairHousingMap'

August 19, 2024

Type Package

Title Create CTCAC/HCD Fair Housing Map

Version 0.1.0

Author Joseph Ahrenholtz

Maintainer josephda@berkeley.edu

Description Code that produces the CTCAC/HCD Fair Housing Map

License What license is it under?

Encoding UTF-8

LazyData true

Imports sf,
sp,
testthat,
tidycensus,
tidyverse

Suggests tigris,
mapview,
testthat (>= 3.0.0),

RoxygenNote 7.2.3

Config/testthat/edition 3

R topics documented:

`all_census_data` *Loads all ACS and decennial Census data into a single data frame.*

Description

Downloads all relevant decennial and ACS data from Census API using tidycensus, then derives percentages and calculates margins of error for derived variables.

Usage

```
all_census_data(
  year = current_year,
  geo = "tract",
  write = FALSE,
  read = !write
)

read_census_data(year = current_year, geo = "tract")

read_acs_data(
  year = current_year,
  geo = "tract",
  testing_handle = FALSE,
  test_name = ""
)
```

Arguments

`year`	designates the map year's filepaths
`geo`	tract or block group
`write`	write the intermediate file
`read`	read an existing intermediate file

Details

To use the Census APIs, sign up for an API key. Add Census API key to .Renviron profile and call it CENSUS_KEY. censusapi will use it by default. Within R, run:
Add key to .Renviron
'Sys.setenv(CENSUS_KEY= 'YOURKEYHERE')'
Reload .Renviron
'readRenviron("~/.Renviron")'
Check to see that the expected key is output in your R console
'Sys.getenv("CENSUS_KEY")'
'tidycensus::census_api_key('YOURKEYHERE', overwrite = TRUE, install = TRUE)'

Value

a data frame

Note

school_distances depends on all_census_data. The 2023 handle is included in order to run 2023 school_distances at 2020 geos, required to implement 3-year rolling averages of education data in the 2025 map.

Source

Census API

tract/bg area: https://mcdc.missouri.edu/applications/geocorr2022.html

Examples

```
all_census_data(read = F) # downloads decennial and acs data for the current map year at the tract level
```

---

create_regions *Create file that distinguishes between regions and rural areas*

Description

'create_regions' evaluates which regions that each county belongs to, then uses 'rural_overlay' to pinpoint rural tracts. 'rural_overlay' merges block centers with the rural_shapefile, and classifies the population of any block with its centroid inside the rural shapefile as rural. It then collapses to tract level, and any tract with over 50 percent population rural is classified as "Rural Areas." 'rural_overlay' is executed within the 'create_regions' function, and is separated for testing convenience only. For creating data, only 'create_regions' is necessary to run.

Usage

```
create_regions(
  year = current_year,
  write = FALSE,
  read = !write,
  testing_handle = FALSE
)

rural_overlay(
  block_points,
  rural_area,
  create_overlay = NULL,
  year = current_year,
  collapse = TRUE
)
```

Arguments

year	designates the map year's filepaths
write	write the intermediate file
block_points	block centroids generated by 'read_block_centers'.
create_overlay	allows for running only the sp::over function separately to reduce debugging time. To use, first assign location_overlay(create_overlay = T) to a variable, then use location_overlay(create_overlay = variable) to run the rest of the function.
collapse	aggregates blocks into tract and assigns tracts to rural or urban designation. Setting this to FALSE allows the user to view the block-level designation.
rural_overlay	rural shapefile.

Value

a data frame

Source

shape_CA_tract created using tigris package: /data-raw/R/generate_census_shapes.R

list of TCAC rural counties: https://www.treasurer.ca.gov/ctcac/Cover-memo.pdf

Examples

```
create_regions(write = T) # computes and writes the region designations file to the intermediate directory
```

---

current_year *Reader functions for zipped files*

Description

Reader functions for zipped files

Usage

```
current_year
```

Arguments

name	file within zip directory
type	Can be 'excel', 'csv', 'tsv', or 'table'

Format

An object of class `numeric` of length 1.

Value

the raw data file

Examples

```
read_zip(acs_variables, year = 2024, type = 'csv')
```

`data_dict_2024` *fairHousingMap data dictionary*

Description

Final 2024 Fair Housing Map data dictionary for descriptions of all variables included in the mapping interface.

Usage

```
?data_dict_2024
```

Details

- fips: Census tract ID.
- fips_bg: Census block group ID.
- region: TCAC regions and Rural Areas designation.
- regionid: Reference region from which regional medians are derived. Is distinct from region field for rural block groups.
- county_name: County name.
- pct_above_200_pov: Percent of population above 200 percent of the federal poverty level.
- home_value: Median home-value of owner occupied units.
- pct_bachelors_plus: Percent of adults with a bachelor's degree or above.
- pct_employed: Percent of adults aged 20-64 who are employed in the civilian labor force or in the armed forces.
- math_prof: Percentage of 4th graders who meet or exceed math proficiency standards.
- read_prof: Percentage of 4th graders who meet or exceed literacy standards.
- grad_rate: Percentage of high school cohort that graduated on time.
- pct_not_frpm: Percent of students not receiving free or reduced-price meals (FRPM).
- env_site_pctl: Regional percentile of averaged site-based environmental hazard indicators.
- pct_above_200_pov_median: Regional median of above 200 percent poverty indicator.
- home_value_median: Regional median of home-value indicator.
- pct_bachelors_plus_median: Regional median of bachelor's+ indicator.
- pct_employed_median: Regional median of employment indicator.
- math_prof_median: Regional median of math proficiency indicator.
- read_prof_median: Regional median of reading proficiency indicator.
- grad_rate_median: Regional median of graduation rate indicator.
- pct_not_frpm_median: Regional median of FRPM indicator.

- env_site_score: Binary score which identifies the bottom 95 percent of site-based environmental hazards. High environmental hazard geographies receive a score of 0, all others receive a score of 1.
- pct_above_200_pov_score: Binary score where 1 denotes the indicator is above the regional median.
- home_value_score: Binary score where 1 denotes the indicator is above the regional median.
- pct_bachelors_plus_score: Binary score where 1 denotes the indicator is above the regional median.
- pct_employed_score: Binary score where 1 denotes the indicator is above the regional median.
- math_prof_score: Binary score where 1 denotes the indicator is above the regional median.
- read_prof_median: Binary score where 1 denotes the indicator is above the regional median.
- grad_rate_score: Binary score where 1 denotes the indicator is above the regional median.
- pct_not_frpm_score: Binary score where 1 denotes the indicator is above the regional median.
- oppscore: Final opportunity score: 9 or 8 = "Highest Resource"; 7 or 6 = "High Resource"; 5 or 4 = "Moderate Resource"; 3 or lower = "Low Resource".
- pct_below_pov: Percent of population below the federal poverty line.
- lq_poc: Location quotient of all people of color relative to the county.
- lq_asian: Location quotient of Asian population relative to the county.
- lq_black: Location quotient of Black population relative to the county.
- lq_hispanic: Location quotient of Hispanic population relative to the county.
- oppcat: Opportunity designation.
- pov_seg_cat: High Poverty and Segregated designation. Defined as least 30 percent of the population falling under the federal poverty line and Location Quotient of higher than 1.25 for Black, Hispanic, Asian, or all people of color in comparison to the county.
- baseline_raceinc0021: Binary flag if tract was a LMI and BIPOC neighborhood in 2000
- baseline_race1321: Binary flag if tract was a BIPOC neighborhood in 2013
- nbrhood_chng: Binary flag if a tract meets the definition of either long-term (2000-2021) or recent neighborhood change (2013-2021).
- trct_raceeth_chng0021: Percentage point change in tract's non-Hispanic white population between 2000 and 2021.
- trct_raceeth_chng1321: Percentage point change in tract's non-Hispanic white population between 2013 and 2021.
- trct_inc_chng0021: Percentage point change in tract's above-moderate-income households between 2000 and 2021.
- trct_inc_chng1321: Percentage point change in tract's above-moderate-income households between 2013 and 2021.
- halfmile_buffer: Binary flag if a tract's population-weighted centroid is within 1/2-mile of the population-weighted centroid of a tract that experienced long-term change (2000-2021).
- raceeth_half0021: Countywide 50 percent threshold for non-Hispanic white tract-level percentage point increase (2000-2021).
- raceeth_half1321: Countywide 50 percent threshold for non-Hispanic white tract-level percentage point increase (2013-2021).
- inc_half0021: Countywide 50 percent threshold for above-moderate-income households tract-level percentage point increase (2000-2021).

- inc_half1321: Countywide 50 percent threshold for above-moderate-income households tract-level percentage point increase (2013-2021).
- rent_quarter1321: Countywide top 25 percent threshold for percent change in median rent (2013-2021).
- trct_pctchng_medrent1321: Percent change in tract's median rent 2013-2021.

`data_dict_2025` *fairHousingMap data dictionary*

Description

Final 2025 Fair Housing Map data dictionary for descriptions of all variables included in the mapping interface.

Usage

```
?data_dict_2025
```

Details

- fips: Census tract ID.
- fips_bg: Census block group ID.
- region: TCAC regions and Rural Areas designation.
- regionid: Reference region from which regional medians are derived. Is distinct from region field for rural block groups.
- county_name: County name.
- pct_above_200_pov: Percent of population above 200 percent of the federal poverty level.
- home_value: Median home-value of owner occupied units.
- pct_bachelors_plus: Percent of adults with a bachelor's degree or above.
- pct_employed: Percent of adults aged 20-64 who are employed in the civilian labor force or in the armed forces.
- math_prof: Percentage of 4th graders who meet or exceed math proficiency standards.
- read_prof: Percentage of 4th graders who meet or exceed literacy standards.
- grad_rate: Percentage of high school cohort that graduated on time.
- pct_not_frpm: Percent of students not receiving free or reduced-price meals (FRPM).
- env_site_pctl: Regional percentile of averaged site-based environmental hazard indicators.
- pct_above_200_pov_median: Regional median of above 200 percent poverty indicator.
- home_value_median: Regional median of home-value indicator.
- pct_bachelors_plus_median: Regional median of bachelor's+ indicator.
- pct_employed_median: Regional median of employment indicator.
- math_prof_median: Regional median of math proficiency indicator.
- read_prof_median: Regional median of reading proficiency indicator.
- grad_rate_median: Regional median of graduation rate indicator.
- pct_not_frpm_median: Regional median of FRPM indicator.

- env_site_score: Binary score which identifies the bottom 95 percent of site-based environmental hazards. High environmental hazard geographies receive a score of 0, all others receive a score of 1.
- pct_above_200_pov_score: Binary score where 1 denotes the indicator is above the regional median.
- home_value_score: Binary score where 1 denotes the indicator is above the regional median.
- pct_bachelors_plus_score: Binary score where 1 denotes the indicator is above the regional median.
- pct_employed_score: Binary score where 1 denotes the indicator is above the regional median.
- math_prof_score: Binary score where 1 denotes the indicator is above the regional median.
- read_prof_median: Binary score where 1 denotes the indicator is above the regional median.
- grad_rate_score: Binary score where 1 denotes the indicator is above the regional median.
- pct_not_frpm_score: Binary score where 1 denotes the indicator is above the regional median.
- oppscore: Final opportunity score: 9 or 8 = "Highest Resource"; 7 or 6 = "High Resource"; 5 or 4 = "Moderate Resource"; 3 or lower = "Low Resource".
- pct_below_pov: Percent of population below the federal poverty line.
- lq_poc: Location quotient of all people of color relative to the county.
- lq_asian: Location quotient of Asian population relative to the county.
- lq_black: Location quotient of Black population relative to the county.
- lq_hispanic: Location quotient of Hispanic population relative to the county.
- oppcat: Opportunity designation.
- pov_seg_cat: High Poverty and Segregated designation. Defined as least 30 percent of the population falling under the federal poverty line and Location Quotient of higher than 1.25 for Black, Hispanic, Asian, or all people of color in comparison to the county.
- trct_raceeth_chng1322: Percentage point change in tract's non-Hispanic white population between 2013 and 2022
- raceeth_half0022: Regionwide 50
- raceeth_half1322: Regionwide 50
- raceeth_quarter1322: Regionwide 75
- baseline_raceinc0022: Binary flag if tract was a LMI and BIPOC neighborhood in 2000
- baseline_race0022: Binary flag if tract was a BIPOC neighborhood in 2000
- baseline_race1322: Binary flag if tract was a BIPOC neighborhood in 2013
- change_race0022: Binary flag if a LMI and BIPOC neighborhood in 2000 experienced an increase in the non-Hispanic white population above the regionwide threshold (2000-2022)
- change_race1322: Binary flag if a BIPOC neighborhood in 2013 experienced an increase in the non-Hispanic white population above the 50
- change_race1321_75pct: Binary flag if a BIPOC neighborhood in 2013 experienced an increase in the non-Hispanic white population above the 75
- exclusion_flag Binary flag if that tract has unreliable data or meets other exclusion parameters (e.g., prisoner population)
- trct_medinc_pctchng_0022: Percent change in tract's household median income between 2000 and 2022
- trct_medinc_pctchng_1322: Percent change in tract's household median income between 2013 and 2022

- medinc_half0022: Regionwide 50
- medinc_half1322: Regionwide 50
- medinc_quarter1322: Regionwide 75
- baseline_income1322: Binary flag if tract was a LMI neighborhood in 2013 (used in rising rents and home value/income gap)
- change_income0022: Binary flag if a LMI and BIPOC neighborhood in 2000 experienced an increase in median household income above the regionwide threshold (2000-2022)
- change_income1322: Binary flag if a LMI and BIPOC neighborhood in 2013 experienced an increase in median household income above the regionwide threshold (2013-2022)
- change_income1321_75pct: Binary flag if a LMI neighborhood in 2013 experienced an increase in median household income above the 75
- trct_pctchng_medrent1322:
- rent_half1322: Regionwide top 50
- medrent_disp1322: Binary flag if an LMI neighborhood in 2013 experienced rent increases above the regionwide threshold (2013-2022)
- income_percentile: Percentile of tract household median income relative to the region (2022)
- hval_percentile: Percentile of tract median home value relative to the region (2022)
- pct_gap: Difference between the tract's household median income percentile (2022) and its median home value percentile (2022)
- hvalinc_gap: Binary flag if the difference between the tract's household median income percentile (2022) and its median home value percentile (2022) is greater than 25 percentage points
- pathway1a: Pathway 1A - Binary flag if a tract meets both racial/ethnic change and economic change between 2000-2022 (50
- halfmile_buffer: Binary flag if a tract's population-weighted centroid is within 1/2-mile of the population-weighted centroid of a tract that meets Pathway 1a
- pathway2: Pathway 2 - Binary flag if a tract that is within 1/2-mile (population-weighted) of a tract that meets Pathway 1A and has rising rents and/or a home value/income gap, and experienced either racial/ethnic or economic change between 2013-2022
- pathway1b: Pathway 1B - Binary flag if a tract meets both racial/ethnic change and economic change between 2013-2022 (75
- nbrhood_chng: Binary flag if a tract meets the definition of neighborhood change (Pathway 1A, Pathway 1B, or Pathway 2)

`filepaths` *Shortcuts to zipped files in the data-raw directory. See '?read_zip'*

Description

Shortcuts to zipped files in the data-raw directory. See '?read_zip'

Usage

```
filepaths(year = current_year)
```

Arguments

year	designates the map year's filepaths

Value

filepaths to raw data are loaded into environment

Note

Year 2023 filepaths are used to process 2023 OM education data at 2020 geographies, which is required for implementation of 3-year rolling education averages.

Examples

```
filepaths(year = 2024)
```

---

final_opp *Returns the final TCAC data frame output*

Description

'final_raw' loads and combines all intermediate files with economic, education, and environmental indicators. 'final_prepare' creates the high poverty and segregated designation and flags unreliable data. 'final_opp' creates the final opportunity scores and designations. 'final_raw' and 'final_prepare' are both inputs into 'final_opp'. Only 'final_opp' is necessary to run for generating new data.

Usage

```
final_opp(year = current_year, write = FALSE, reduced = TRUE, cog = FALSE)

final_raw(year = current_year, geo = "tract")

final_prepare(year = current_year, geo = "tract", .data = NULL)
```

Arguments

year	designates the map year's filepaths
write	write the final output
cog	logical to write COG referenced shapefile for HCD AFFH Data Viewer. Is tract-only. Counties within COG are COG-referenced. Counties outside of COGS are county-referenced.
geo	allows for opportunity to be assessed at the tract level in urban areas and block group in rural areas
output	'reduced' returns only variables necessary for the map. If 'full', include all intermediate variables.

Value

a data frame

Note

2025 implements a 3-year rolling average of education indicators. The change is enacted in the final_raw function.

Examples

```
final_opp(year = 2024, write = TRUE) # writes the final output
```

---

final_opp_public *Generates public summary files*

Description

Loads final opportunity and neighborhood change data and creates excel workbook summary files, shapefiles, and data dictionaries

Usage

```
final_opp_public(year = current_year, write = FALSE, change = FALSE)
```

Arguments

year	designates the map year's filepaths
write	whether to write new opportunity summary files
change	whether to write new change summary files

Examples

```
final_opp_public(year = 2024, write = TRUE, change = TRUE)
```

---

hello *Hello, World!*

Description

Prints 'Hello, world!'.

Usage

```
hello()
```

Examples

```
hello()
```

opp_lihtc_devels	*Plot 9 Preservation Database extraction share by CHPC 2024 CTCAC/HCD Opportunity Map OBI Categorical segregation: https://berkeley.app.box.com/file/1298215376451 opp_lihtc_devels() Plot 9 Cutoff year is 2019 for 9 with application years 2015 or newer for 9 the spatial join of lihtc developments to neighborhood categories. plot_lihtc_devels() reads the output of the former, summarizes and plots the data, and the figures are saved as .tiff files in the products/charts directory. The code below was adjusted in March, 2024 per HCD request to treat High-Poverty & Segregated as an exclusive category, though in the 2024 map it was shifted to an overlay. This will likely change in future iterations as the research partners identify an alternative way to present the data that is more aligned with the current mapping methodology. Additionally, this code was written in parallel to CHPC updating their geocoding standardization process. Once this process is finalized, the input files or fields may be adjusted. Currently, CHPC is sharing one file for locations and a second file with project details. The files can be difficult to join because of duplicate id's, project names, app id's, and addresses. OBI should request that they share a single file to avoid these issues. The two function calls currently do not include parameters. If the charts do not change significantly in 2025 than parameters can be written to specify the year. If the 2025 charts change significantly than the 2024 code may need to be archived and a new script written. The charts produced in 2022 are not reproducible because many projects needed to be manually located, which spurred efforts in 2023 between CHPC and OBI to standardize and document the geocoding process . There were no charts made using the 2023 map. opp_lihtc_devels() # call first plot_lihtc_devels() # call second*

read_educ_pov	*Imports education data and writes the relevant variables to the intermediate directory*

Description

Imports education data and writes the relevant variables to the intermediate directory

Usage

```
read_educ_pov(year = current_year)
```

Arguments

year	designates the appropriate filepaths
write	write the intermediate file
read	read an existing intermediate file

Details

Reading data is weighted by total enrollment of schools that return 4th-grade test scores. FRPM data is limited to and weighted by enrollment in schools that serve elementary school students. Grad data is weighted by high school enrollment.

Value

a data frame

Source

School data: https://www.cde.ca.gov/ds/si/ds/pubschls.asp

Test data: https://caaspp-elpac.ets.org/caaspp/ResearchFileListSB?

FRPM: http://www.cde.ca.gov/ds/sd/sd/filessp.asp

Cohort Grad data: https://www.cde.ca.gov/ds/ad/filesacgr.asp

Examples

```
graduation_rates() # reads an existing intermediate file at the default year
```

`read_neighborhood_change`

Import Neighborhood Change Data

Description

Imports neighborhood change data generated by CHPC identifying non-rural tracts that have experienced both long-term (since 2000) and recent change (since 2013) racial/ethnic and economic change.

Usage

```
read_neighborhood_change(year = current_year)
```

Arguments

`year`	designates the map year's filepaths

Examples

```
neighborhood_change(year = 2024) # imports neighborhood change output and reduces to necessary variables
```

read_tract_centers *Reads pop-weighted centroids of Census tracts, block groups, and blocks*

Description

Reads pop-weighted centroids of Census tracts, block groups, and blocks

Usage

```
read_tract_centers(as_shape = FALSE, year = current_year)
```

Arguments

as_shape	returns shapefile
year	designates the map year's filepaths

Value

a data frame

Source

tract/bg centroids: https://www.census.gov/geographies/reference-files/time-series/geo/centers-population.html

block centroids: https://mcdc.missouri.edu/applications/geocorr2022.html

Examples

```
read_tract_centers(as_shape = T) # returns tract centroids as shapefile based on the default year
```

school_distances *Creates tract/block group education scores*

Description

Creates tract/block group education scores

Usage

```
school_distances(
  year = current_year,
  geo = "tract",
  write = FALSE,
  read = !write
)
```

Arguments

year	designates the appropriate filepaths
geo	tract or bg
write	write the intermediate file
read	read an existing intermediate file

Details

Finds school distance to tract or block group centroids, and averages the reading and math scores, frpm, and graduation rates of the three nearest schools. 4th grade and FRPM are weighted by the enrollment of schools that serve 4th-graders, and graduation rates are weighted by cohort size.

Value

a data frame

Examples

```
school_distances(year = 2024, geo = 'tract', write = TRUE, read = FALSE) # writes a new file to the intermediate
```

---

shape_rural *Create rural designation shapefile*

Description

Imports USDA shapefile for areas ineligible for rural designation, then shrinks to only California data and subtracts them from the complete California shapefile to get a rural areas shapefile.

Usage

```
shape_rural(year = current_year, write = FALSE, testing_handle = FALSE)
```

Arguments

year	designates the map year's filepaths
write	whether to write the intermediate shapefile

Value

a shapefile

Note

as of 06/2024, eligibility file is downloading as a geodatabase with null values in the state name field. The "_CA" suffixed file was pre-processed in ArcGIS. May consider reaching out to someone at USDS to see why fields are null before publishing.

Source

USDA areas ineligible for rural development housing programs: https://www.sc.egov.usda.gov/data/data_files.html

Changes in 2019 detailed: https://www.rd.usda.gov/files/CA-SFH-NoticeRuralAreaReview-Final-4.16.18.pdf

Examples

```
shape_rural(write = T) # computes and writes the rural shapefile to the intermediate directory
```

---

tribal_overlap *Create tribal land flag*

Description

Imports tribal lands under of the control of federally-recognized tribes, computes intersection with Census tracts, and flags any tract where at least 25 percent of the geography's land area is within federally-recognized tribal lands. In final_data.R, High-Poverty & Segregated is not assessed in tracts where the tribal land flag is raised.

Usage

```
tribal_overlap(year = current_year)
```

Arguments

year	designates the map year's filepaths

Value

a dataframe

Examples

```
tribal_overlap(year = 2024) # loads tracts with flag for tribal land
```

xwalk_ces *Create site-based environmental hazards variable*

Description

Imports CalEnviroScreen 4.0 site-based measurements (cleanup sites, hazardous waste, groundwater threats, solid waste) from the final 2023 TCAC file, creates a binary score for tracts in the bottom 5 crosswalks to 2020 tracts using an overlay method of >= 5 until OEHHA updates to 2020 boundaries. The overlay approach was chosen over a weighted allocation (e.g. by area or population) because the site-based measures are already interpolated by OEHHA from points to tracts, and the research partners wanted to avoid re-interpolating already interpolated data.

Usage

```
xwalk_ces(
  year = current_year,
  write = FALSE,
  read = !write,
  testing_handle = FALSE
)
```

Arguments

year	designates the appropriate filepaths
write	write the intermediate file
read	read an existing intermediate file

Value

a data frame

Examples

```
xwalk_ces(year = 2024, write = TRUE) # writes a new file to the intermediate directory
```

Index