# Redlining Data

Author: Isaac Rand

Affiliation(s): HEROP, University of Chicago

Last updated: 4 May 2021

## Data Columns

* GEOID: 11 digit FIPS identifier.
* proportionA: The proportion of the land area of the tract which was categorized as “Best,” “First Grade,” or simply “A.” This was the most favorable category assigned by HOLC.
* proportionB: The proportion of the land area of the tract which was categorized as “Still Desirable,” “Second Grade,” or simply “B.” This was the second most favorable category assigned by HOLC.
* proportionC: The proportion of the land area of the tract which was categorized as “Definitely Declining,” “Third Grade,” or simply “C.” This was the second least favorable category assigned by HOLC, and residents of areas with this categorization were routinely denied financial and other services.
* proportionD: The proportion of the land area of the tract which was categorized as “Hazardous,” “Fourth Grade,” or simply “D.” This was the least favorable category assigned by HOLC, and residents of areas with this categorization were routinely denied financial and other services.
* proportionCorD: The proportion of the land area of the tract which was categorized as C or D. Since residents of areas with either of these categorizations were routinely denied financial and other services, it may make sense to consider the categories together.
* primary\_grade\_4levels: This variable shows which of the four redlining categories (A, B, C, or D) was assigned to the most area within the tract.
* primary\_grade\_3levels: This variable also reports which of the redlining categories was assigned to the most area in the tract, but combines C and D into one category CD. Since residents in areas assigned to both categories were routinely discriminated against, it may make sense to consider them together.

## Map Variables

**Redlining Categorization (4 level):** Redlining maps of US cities were originally created by the Home Owners’ Loan Corporation (HOLC) in the 1930s. Neighborhoods were assigned to one of four categories: A - “Best,” B - “Still Desirable,” C - “Definitely Declining,” or D - “Hazardous.” This variable reports which of the four categories was assigned to the most area within each tract, for all tracts in which over 50% of the tract area was categorized. Residents in areas marked C or D were frequently denied financial and other services, and the legacy of discrimination against these areas persist materially today.

**Redlining Categorization (3 level):** Redlining maps of US cities were originally created by the Home Owners’ Loan Corporation (HOLC) in the 1930s. Neighborhoods were assigned to one of four categories: A - “Best,” B - “Still Desirable,” C - “Definitely Declining,” or D - “Hazardous.” This variable reports which of the categories was assigned to the most area within each tract, for all tracts in which over 50% of the tract area was categorized, but combines the C and D categories into one CD category. Residents in areas marked either C or D were frequently denied financial and other services, and the legacy of discrimination against these areas persist materially today.

## Sources and Years

* HOLC Boundaries, Robert K. Nelson, LaDale Winling, Richard Marciano, Nathan Connolly, et al., “Mapping Inequality,” American Panorama, ed. Robert K. Nelson and Edward L. Ayers, accessed May 4, 2022, [Mapping Inequality (richmond.edu)](https://dsl.richmond.edu/panorama/redlining/#loc=5/39.1/-94.58&text=downloads)[/](https://dsl.richmond.edu/panorama/redlining/), late 1930’s

## Data Processing

* Data processing was done in R. For each tract which spatially intersected any HOLC boundaries, a proportion was calculated reporting how much of that tract was covered by each HOLC category (A, B, C, or D). Whichever of these proportions was highest is reported in “Redlining Categorization (4 levels).” “Redlining Categorization (3 levels)” is calculated the same way, but with the proportions assigned to C or D counted together. This way, if a tract is 40% B, 20% C, 30% D, it will be marked as primarily CD in “Redlining Categorization (3 levels)” since 50% of the tract is C or D, and primarily B in “Redlining Categorization (4 levels).”

## Original Spatial Scale

* The HOLC boundaries are polygons which are not nested into any conventional spatial scale. The average boundary covers about 0.54 square miles, with a standard deviation of about 0.93 square miles. They were agglomerated to census tracts as described in data processing.

## Limitations

* Tracts do not perfectly align with HOLC boundaries, and not all tracts are completely covered by HOLC categories (some areas, especially non-residential ones, are uncategorized). The data here only includes those tracts in which more than 50% of the tract area was categorized by HOLC.

## Organizational and Data Source Links

## Data Source - <https://dsl.richmond.edu/panorama/redlining/#loc=5/39.1/-94.58&text=downloads>[/](https://dsl.richmond.edu/panorama/redlining/)

## Additional Notes

Source R code for data processing is available here:

[geospatial-legacy-redlining/find\_tracts\_proportion\_redlined.Rmd at main · GeoDaCenter/geospatial-legacy-redlining (github.com)](https://github.com/GeoDaCenter/geospatial-legacy-redlining/blob/main/public/geojson/Redlining%20at%20Tract%20Level%20Metadata/find_tracts_proportion_redlined.Rmd)