This data was obtained from the Consumer Financial Protection Bureau on June 7, 2023 and was heavily massaged by Dr. Wynne to get in a format suitable for class work.

Original Data Source:

<https://ffiec.cfpb.gov/data-browser/data/2021?category=states> and filter to County = 06059 ORANGE COUNTY – CALIFORNIA

The “institution” variable is originally listed as a Legal Entity Identifier (LEI), which can be matched with an institution name using the Global Legal Entity Identifier Foundation (GLEIF) search at <https://search.gleif.org/#/search/>

The “city” variable is not contained in the original data and was found by looking up the INTPTLAT and INTPTLON values from <https://data-ocpw.opendata.arcgis.com/datasets/OCPW::ocacs-2020-demographic-characteristics-for-census-tracts/explore?showTable=true>

Variables in this dataset:

* institution (chr): the financial institution that received the loan application. Note that most of the values here are "Other" because Dr. Wynne got tired of manually searching the Legal Entity Identifiers for institutions with only a small number of applications.
* census\_tract (chr): the census tract corresponding to the address listed on the application. Note that this should be a series of 6 digits. Please see the map for boundaries; note that the tract numbers on the map use the same sequence of digits but a slightly different format.
* ethnicity (chr): the ethnicity of the applicant(s). Note that "Joint" means that one or more applicants identified as Hispanic or Latino and one or more applicants identified as Not Hispanic or Latino.
* race (chr): the race of the applicant(s). Note that "Joint" means that one or more applicants identified as White and one or more applicants identified as a different race.
* sex (chr): the sex of the applicant(s). Note that "Joint" means that one or more applicants identified as Female and one or more applicants identified as Male.
* residency\_type (chr): the intended use of the property (Primary Residence, Secondary Residence, or Investment Property)
* loan\_amount (dbl): the requested loan amount, in US dollars (USD)
* property\_value (dbl): the appraised value of the property, in USD
* interest\_rate (dbl): the interest rate of the loan, in percentage points
* total\_loan\_costs (dbl): the amount of total loan costs (other than principal and interest), in USD
* income (dbl): the gross annual income relied on in making the application decision, in thousands of USD
* applicant\_over\_62 (lgl): whether any applicant was over the age of 62
* tract\_population (dbl): the population of the census tract corresponding to the property address
* tract\_minority\_population\_percent (dbl): the percentage of the census tract that identifies as a racial or ethnic minority
* tract\_to\_msa\_income\_percentage (dbl): the ratio of census tract median family income to the median family income for the metropolitan statistical area, in percent (i.e., 100 = exactly at median)
* tract\_housing\_median\_age (dbl): the median age of homes in the census tract, in years
* city (chr): the city within which the geographic center of the census tract is located. In most, but not all cases, this will correspond to the city the property is located in.
* action (chr): whether the loan was Approved or Not approved.