Codebook

Codebook for working_dataset.R.
177 Variables, 4386 Rows (district-years)

Variable	Description
city	City name
district	District name
district_name	Alternative name of the district (e.g. district = Altstadt, district_name = IV)
subdistricts	Name of the subdistrict nested in the district
iso	3-digit country ISO code
year	Observation year
source	Original name of the datasource
dropbox_link	Link to original source on Dropbox
area.size.ha	Total area in hectares
residential.buildings	Number of residential buildings
population	Total population in the district
households	Total number of households
housing.units	Number of housing units
total.rooms.res.plus. non_residential	Total number of rooms (residential + non-residential)
heatable.rooms	Number of rooms that can be heated
non_heatable	Number of rooms without heating
kitchen	Number of kitchens
kitchen_2	Correction indicator for unit_rooms variable, 1 = kitchen is counted as a room, 0 = kitchen is not counted as a room
housing.rooms	Total number of housing rooms in housing units
occupants.per. 100.housing.rooms	Number of occupants per 100 housing rooms
vacant	Number of vacant housing units
rent.units	Number of tenant units
average.rent	Average rent [1]
currency	Currency used in rent and income variables
average.income.per.unit	Average income per housing unit (city level)
income.per.district	Average income per housing unit (district level)
owner_occupied.units	Number of owner-occupied housing units
dwelling_unit01– dwelling_unit50plus	Distribution of dwellings by number of housing units in a building (e.g. dwelling_unit01 = Number of dwellings with one housing unit in building/single family dwellings)
unit_rooms00– unit_rooms_more	Distribution of housing units by number of rooms (unit_rooms00 = Number of housing units with no standard housing rooms)

[1] Directly surveyed, if NA calculated as a weighted average by room-segment shares

Variable	Description
occ_rooms00– occ_rooms_more	Distribution of occupants by number of rooms (occ_rooms00 = Number of occupants in housing units with no standard housing rooms)
rent_rooms00– rent_rooms10	Average rent per housing unit by the number of rooms (rent_rooms00 = Average rent for a housing unit with no standard housing rooms)
units_stories.-01– stories_more	Distribution of housing units by number of stories (units_stories.-01 = Number of housing units in the cellar)
city_year	Combined city-year identifier
persons_building	Persons per building
persons_unit	Population per unit
rooms_total	Total number of housing rooms [2]
persons_room	Persons per room
persons_ha	Persons per hectare
dens_rooms00– dens_rooms10	Calculated per room densities [3]
density	Persons per hectare
own	Share of owner-occupied housing
own1	Share of owner-occupied housing (calculated through the share of rent units)
rent	Share of rented housing
label	Name of the city (label used in creating figures)
households_unit	Households per unit
units_building	Housing units per building
pop_household	Persons per household
rooms_unit	Rooms per unit
total_units	Total number of housing units [4]
share_unit0– share_unit6more	Share of housing units by number of housing rooms
dwelling_unit12– dwelling_unit_more	Grouped version of the dwelling_unit01–dwelling_unit50plus variables into higher-order categories.
units_ha	Housing units per hectare
buildings_ha	Buildings per hectare
rooms_ha	Rooms per hectare
city_pop	Total population of the city in observation year
pop1800, pop1900	City population estimates in 1800 and 1900 [5]
hhsize	Average household size

[2]Calculated from original data if housing.rooms variable was NA, sumproduct of unit_roomsXX variables*number of rooms

[3]e.g. occ_room01/(unit_rooms01*1)

[4]Calculated from original data, rowsum of unit_roomsXX variables

[5]Based on data provided by Bairoch, P. (1988). Cities and economic development: from the dawn of history to the present. University of Chicago Press.; Chandler, T. Four Thousand Years of Urban Growth: An Historical Census (The Edwin Mellen Press, 1987). and Bolt, Jutta; Van Zanden, Jan Luiten, 2024, "Maddison Project Database 2023", https://doi.org/10.34894/INZBF2, DataverseNL, V1

Variable	Description
GDP/capita****,	GDP per capita in 1800, 1850, 1900 [5][6]
country.pop1800/1900	National population in corresponding years [6]
city_spurt	City population growth rate between 1800 and 1900
country_spurt	Country population growth rate between 1800 and 1900
gdp_spurt	GDP per capita growth rate per country between 1800 and 1900
city_age	Age of the city [6][7]
l* prefixes	
(e.g. laverage.rent)	Log-transformed value of the original variable (Log-transformed average rent)
mpop	Average persons per unit
_scaled suffixes	
(e.g. persons_unit_scaled)	Scaled (standardized) versions of key variables
rooms_adjusted	Adjusted total number of housing rooms (harmonized across censuses using the kitchen2 variable)
rooms00_adjusted–rooms_more_adjusted	Adjusted room categories (harmonized across censuses using the kitchen2 variable)

[6] In 2011 USD

[7] Calculated by subtracting founding year (year in which a city counted 5000 inhabitants for the first time) - current year (2024))