The dataset contains the following columns:

PARCELNO: unique identifier for each property. About 1% appear multiple times.

SALE\_PRC: sale price ($)

LND\_SQFOOT: land area (square feet)

TOTLVGAREA: floor area (square feet)

SPECFEATVAL: value of special features (e.g., swimming pools) ($)

RAIL\_DIST: distance to the nearest rail line (an indicator of noise) (feet)

OCEAN\_DIST: distance to the ocean (feet)

WATER\_DIST: distance to the nearest body of water (feet)

CNTR\_DIST: distance to the central business district (feet)

SUBCNTR\_DI: distance to the nearest subcenter (feet)

HWY\_DIST: distance to the nearest highway (an indicator of noise) (feet)

age: age of the structure

avno60plus: dummy variable for airplane noise exceeding an acceptable level

structure\_quality: quality of the structure

month\_sold: sale month in 2023 (1 = jan)

LATITUDE

LONGITUDE