Project 2: Real-Estate Tycoon

Data Dictionary

For this project:

- The dataset has **1883** observations in the county where the REIT operates.
- Each observation is for the transaction of one property only.
- Each transaction was between $200,000 and $800,000.

We have the following features:

Target variable

- `'tx_price'` - Transaction price in USD

Public records for the property

- `'tx_year'` - Year the transaction took place
- `'property_tax'` - Monthly property tax
- `'insurance'` - Cost of monthly homeowner's insurance

Property characteristics

- `'beds'` - Number of bedrooms
- `'baths'` - Number of bathrooms
- `'sqft'` - Total floor area in squared feet
- `'lot_size'` - Total outside area in squared feet
- `'year_built'` - Year property was built
- `'basement'` - Does the property have a basement?

Location convenience scores

- `'restaurants'` - Number of restaurants within 1 mile
- `'groceries'` - Number of grocery stores within 1 mile
- `'nightlife'` - Number of nightlife venues within 1 mile
- `'cafes'` - Number of cafes within 1 mile
- `'shopping'` - Number of stores within 1 mile
- `'arts_entertainment'` - Number of arts and entertainment venues within 1 mile
- `'beauty_spas'` - Number of beauty and spa locations within 1 mile
- `'active_life'` - Number of gyms, yoga studios, and sports venues within 1 mile

Neighborhood demographics

- `'median_age'` - Median age of the neighborhood
- `'married'` - Percent of neighborhood who are married
- `'college_grad'` - Percent of neighborhood who graduated college

Schools

- `'num_schools'` - Number of public schools within district
- `'median_school'` - Median score of the public schools within district, on the range 1 - 10