

# AE 06: Prediction for MLR

## Houses in Levittown

Sep 26, 2022

### ! Important

The AE is due on GitHub by Thursday, September 29 at 11:59pm.

```
library(tidyverse)
library(tidymodels)
library(knitr)
```

The data set contains the sales price and characteristics of 85 homes in Levittown, NY that sold between June 2010 and May 2011. Levittown was built right after WWII and was the first planned suburban community built using mass production techniques.

```
levittown <- read_csv("data/homeprices.csv")
```

The variables used in this analysis are

- **bedrooms:** Number of bedrooms
- **bathrooms:** Number of bathrooms
- **living\_area:** Total living area of the house (in square feet)
- **lot\_size:** Total area of the lot (in square feet)
- **year\_built:** Year the house was built
- **property\_tax:** Annual property taxes (in USD)
- **sale\_price:** Sales price (in USD)

The goal of the analysis is to use the characteristics of a house to understand variability in the sales price.

## Linear model

```
price_fit <- linear_reg() |>
  set_engine("lm") |>
  fit(sale_price ~ bedrooms + bathrooms + living_area + lot_size +
      year_built + property_tax, data = levittown)

tidy(price_fit) |>
  kable(digits = 3)
```

term	estimate	std.error	statistic	p.value
(Intercept)	-7148818.957	3820093.694	-1.871	0.065
bedrooms	-12291.011	9346.727	-1.315	0.192
bathrooms	51699.236	13094.170	3.948	0.000
living_area	65.903	15.979	4.124	0.000
lot_size	-0.897	4.194	-0.214	0.831
year_built	3760.898	1962.504	1.916	0.059
property_tax	1.476	2.832	0.521	0.604

## Prediction

What is the predicted sale price for an individual house in Levittown, NY with 4 bedrooms, 2 bathrooms, 1,800 square feet of living area, 6,000 square foot lot size, built in 1947 with \$7,403 in property taxes?

Report the predicted value and appropriate interval.

### Note

Fill in the code, then make `#| eval: true` before rendering the document.

```
# create tibble for new observation
new_house <- tibble(
  bedrooms = 4,
  bathrooms = 2,
  living_area = 1800,
  lot_size = 6000,
  year_built = 1947,
  property_tax = 7403
)
```

```
# prediction + interval  
predict(price_fit, new_house, type = "pred_int", level = .95)
```

- Interpret the interval in the context of the data.

We are 95% confident that the true sales price of a house in Levittown, NY with 4 bedrooms, 2 bathrooms, 1800 square feet of living area, 6000 square foot lot size, built in 1947 with /\$7,403 in annual property taxes lies on the interval (255576, 448532).

### ! Important

To submit the AE:

- Render the document to produce the PDF with all of your work from today's class.
- Push all your work to your `ae-06-` repo on GitHub. (You do not submit AEs on Gradescope).