Assignment 1: Implement your own Linear Regression

Get the data here

Objective - **Perform linear regression to make a model to find house prices.**

1. Load and preprocess the dataset.
2. Perform exploratory data analysis.
3. Implement linear regression using the matrix-based method taught in class.
4. Evaluate the performance of your regression model.

Guidelines

- **Libraries you can use**: `numpy`, `pandas`, `matplotlib`, `seaborn` but **NOT** any ML or stat library
- Follow the matrix-based regression method taught in class.

Evaluate the Model

- Calculate performance metrics, including:
  - Mean Squared Error (MSE)
  - Residual sum of squares (RSS)
- Interpret your model's performance

Submission

- Submit a Jupyter Notebook with your code and explanations.
- Include visualizations, outputs, and any observations.

Bonus Task (Optional)

1. Experiment with standardising your independent variables. Standardization transforms input features with a mean of 0 and a standard deviation of 1.

For a feature x:

$$z = \frac{x - \mu}{\sigma}$$

Where:

- z: Standardized value.
- x: Original feature value.
- μ: Mean of the feature.
- σ: Standard deviation of the feature.

2. Extend your implementation to include polynomial regression by creating polynomial features.

A general polynomial regression equation:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \ldots + \beta_n x^n + \epsilon$$

Submission Deadline

- Date: 8 JUNE 2025 EOD
- Late submissions will incur a penalty.