



# Rensselaer

why not change the world?®

## Lab2: Linear Regression

**Ahmed Eleish**

**February 6th, 2026**

**ITWS/CSCI/MGMT - 4600/6600 BCBP - 4600**

Tetherless World Constellation  
Rensselaer Polytechnic Institute



# Lab 2

Files:

<https://rpi.box.com/s/qqdy38yjqa1km1om7no5crxm8top696>



# Exercise 1: linear models

- Your exercise: Examine the influence of various variables on property price
- Dataset: NY-House-Dataset.csv



## Using the code in NYHousing\_Linear\_Regression.R do the following:

- Fit **3** linear models with Price as the response variable and with combinations of PropertySqFt, Beds, and Bath as predictors.
- Do any data cleaning (outliers, etc.) to find the best performing models. Remember that you may need to transform variables for a better fit.
- For each model:
  - Print model summary stats.
  - Plot the most significant variable vs Price with the best fit line.
  - Plot a scatter plot of the residuals for the model.
- Compare the 3 models and decide which you believe is most useful. You don't need to tell me which one.



Please upload to your github repository or LMS:

1. All your code in a \*.R or \*.MD file



Thanks!  
Have a great weekend!