Programming Assignment
CL 202 Spring 2021
Due Date: April 21, 2021

This assignment must be solved using a software and concerns the real estate industry. You can do calculations from first principles but all statistical software have in-built commands. Please use these commands. To enable use of an autograder the final answer to the question posed should be assigned to a specific variable name provided in the question itself.

From existing data, we would like to build a regression model that relates the size of an apartment and its ready reckoner value to the selling price of the apartment. The Table below gives data for 20 houses:

z_1 Total dwelling size ($100\ ft^2$)	z_2 Assessed value ($1000)	Y Selling price ($1000)
15.31	57.3	74.8
15.20	63.8	74.0
16.25	65.4	72.9
14.33	57.0	70.0
14.57	63.8	74.9
17.33	63.2	76.0
14.48	60.2	72.0
14.91	57.7	73.5
15.25	56.4	74.5
13.89	55.6	73.5
15.18	62.6	71.5
14.44	63.4	71.0
14.87	60.2	78.9
18.63	67.2	86.5
15.20	57.1	68.0
25.76	89.6	102.0
19.05	68.6	84.0
15.37	60.1	69.0
18.06	66.3	88.0
16.35	65.8	76.0

a) Find correlation coefficient between i) Selling Price and Dwelling size (call this variable *corr_coeff_Y_Z1*), and ii) Selling Price and Assessed Value (*corr_coeff_Y_Z2*)

b) Find a linear model between Y = a*Z1+ b*Z2 + c. Please make sure your code assigns results to *a, b, c*

c) Provide a 95% confidence interval for the three parameters. Please assign the lower value bounds to *ci_l_a, ci_l_b, ci_l_c* and upper bounds to *ci_u_a, ci_u_b, ci_u_c*.

d) Provide a 95% prediction interval for the selling price of an apartment of size 12,00 ft2 and an assessed value of $60,000. Please assign the lower value bound to *pi_l_Y* and upper bound to *pi_u_Y*.

e) Find mean and variannce of residuals. Please assign your answers to variables *mean_resid, var_resid*

f) Find R^2 of the fit. Please assign answer to *Rsq*