STA 5207 Assignment 1

Due Friday September 10

You do not have to turn in R (or SAS) output or code.

A commercial real estate company evaluates vacancy rates, square footage, rental rates, and operating expenses for commercial properties in a large metropolitan area in order to provide clients with quantitative information upon which to make rental decisions. The data from 81 properties consists of rental rates (*y*) in percent, age (*x*1) in years, operating expenses and taxes (*x*2), vacancy rates (*x*3) in proportions, and cost (*x*4) in dollars. The data set is “properties.txt” in Canvas under Files->Data Sets and is in the Data Sets folder in SAS Studio. The file in Canvas has headings but the one in SAS Studio does not.

1. (5 points) Give the multiple linear regression model using all 4 predictors, and the assumptions about the error terms.
   1. Model: yi = β0 + β1xi1 + β2xi2 + β3xi3 + β4xi4 + εi
   2. Assumptions: The error terms εi are independent and identically distributed normal random variables with mean 0 and variance σ2.
2. (5 points) Give the interpretation of β4 in the model.
   1. If the cost of the property increases by $1, while the age, operating expenses and taxes, and vacancy rate stay the same, the mean rental rate changes by β4 units.
3. (10 points) Give the estimated regression equation using all 4 predictors. What is the interpretation of
   1. If a properties age increases by 1 year, while the operating expenses and taxes, vacancy rate, and cost stay the same, we predict that the mean rental rate will decrease by 0.142 units.
4. (10 points) Test for an overall regression relationship at the 5% significance level. Give the hypotheses, test statistic, P-value, and conclusion.

H1: At least one of them is not 0

* 1. *F* = 26.76 P-value = 7.272e-14
  2. Reject H0. There is an overall regression relationship between at least one of the predictors and the response.

1. (15 points) Give individual 95% confidence intervals for the slope parameters.

2.5 % 97.5 %

(Intercept) 1.104949e+01 1.335169e+01

Age -1.845411e-01 -9.952615e-02

Expenses 1.561979e-01 4.078352e-01

Vacancy Rates -1.545232e+00 2.783919e+00

Cost 5.166283e-06 1.068232e-05

1. (15 points) Conduct a t-test at the 5% significance level for each slope parameter, giving hypotheses, test statistics, P-values, and conclusions. What do the conclusions imply about the linear relationships between the predictors and the rental rate?

Age: H0: β1 = 0 H1: β1 ≠ 0

t = -6.655 P-value < 0.0001

Reject H0. Age has a significant linear relationship with rental rate after accounting for expenses and taxes, vacancy rates, and cost.

Expenses: H0: β2 = 0 H1: β2 ≠ 0

t = 4.464 P-value < 0.0001

Reject H0. Taxes have a significant linear relationship with rental rate after accounting for age, vacancy rates, and cost.

Vacancy Rates: H0: β3 = 0 H1: β3 ≠ 0

t = 0.570 P-value = 0.57

Do not reject H0. Vacancy rates do not have a significant linear relationship with rental rate after accounting for age, expenses and taxes, and cost.

Cost: H0: β4 = 0 H1: β4 ≠ 0

t = 5.722 P-value < 0.0001

Reject H0. Cost has a significant linear relationship with rental rate after accounting for age, expenses and taxes, and vacancy rates.

1. (10 points) Give the value and interpretation of R2.
   1. R­2 = 0.5847. 58.47% of the total variation in rental rates can be explained by the 4 predictors.
2. (5 points) Give a 95% confidence interval for the mean rental rate of 5 year old properties with 4.1 tax rate, 0.16 vacancy rate, and cost of $100,000 square feet.
   1. (12.91148, 14.16494)
3. (5 points) Give a 95% prediction interval for a single property with the predictor values given in question 8.
   1. (11.18877, 15.88765)
4. (20 points) Perform simple linear regression using rental rate as the response and:
5. age as the only predictor. Does this t-test match the result of the test in question 6?
   1. t = -2.23, P-value = 0.0242
   2. Yes, it still matches for significance level of 0.05.
6. expenses as the only predictor. Does this t-test match the result of the test in question 6?
   1. t = 4.04, P-value = 0.000123
   2. Yes, it matches.
7. vacancy rate as the only predictor. Does this t-test match the result of the test in question 6?
   1. t = 0.593, P-value = 0.555
   2. Yes, it matches.
8. cost as the only predictor. Does this t-test match the result of the test in question 6?
   1. t = 5.632, P-value < 0.0001
   2. Yes, it matches.