MA 542: REGRESSION ANALYSIS

QUIZ - 5 **Name** Key

The following are the R-outputs for COMMERCIAL PROPERTIES example. (Y-rental rates, X1-age, X2-expences and taxes,X3-vacancy rates, X4-squre footage). The first one is the summary output of the model for all the predictors and the second one is the summary output of the model for predictors X1-age, X2-expences and taxes and X4-squre footage.

```
Call:
lm(formula = Y ~ X1 + X2 + X3 + X4)

Residuals:
    Min      1Q  Median      3Q     Max
-3.1872 -0.5911 -0.0910  0.5579  2.9441

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.220e+01  5.780e-01  21.110  < 2e-16 ***
X1          -1.420e-01  2.134e-02  -6.655 3.89e-09 ***
X2           2.820e-01  6.317e-02   4.464 2.75e-05 ***
X3           6.193e-01  1.087e+00   0.570     0.57
X4           7.924e-06  1.385e-06   5.722 1.98e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.137 on 76 degrees of freedom
Multiple R-squared:  0.5847,   Adjusted R-squared:  0.5629
F-statistic: 26.76 on 4 and 76 DF,  p-value: 7.272e-14
```

> 0.05

```
Call:
lm(formula = Y ~ X1 + X2 + X4)

Residuals:
    Min      1Q  Median      3Q     Max
-3.0620 -0.6437 -0.1013  0.5672  2.9583

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.237e+01  4.928e-01  25.100  < 2e-16 ***
X1          -1.442e-01  2.092e-02  -6.891 1.33e-09 ***
X2           2.672e-01  5.729e-02   4.663 1.29e-05 ***
X4           8.178e-06  1.305e-06   6.265 1.97e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.132 on 77 degrees of freedom
Multiple R-squared:  0.583,    Adjusted R-squared:  0.5667
F-statistic: 35.88 on 3 and 77 DF,  p-value: 1.295e-14
```

a) Which model do you choose for the data given? Why? (Give two reasons for your conclusion)

Model-2 is chosen, because

1. predictor X_3 (vacancy rates) is not significant in model-I

and

2. Adjusted R^2 value is larger for the second modul than that for the first model.

b) Perform a four step hypotheses test to test whether there is a regression relationship between rental rates and **four** predictors.

Hypotheses

1) $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ vs H_1: not H_0

2) Test Statistic:

$F = 26.76$ (from the output)

3) p-value $= 7.272 \times 10^{-14}$ (from the output)

4) conclusion:

Since p-value < 0.05 (α), H_0 is rejected.

So there is a regression relationship between the rental rates and predictors age, expences and taxes, vacancy rates and Square footage.