

MTH 4230 Lab 1 **Answer Sheet**

Due Wed., Feb. 5

1 Part A: Simple Linear Regression

1.1 Murder Rates Data Set

1. NA
2. NA (*don't* print the scatterplot).
3. NA
4. NA (*don't* print the scatterplot).
5. Give the following values:

The estimate of β_0 is $b_0 =$ _____

The estimate of β_1 is $b_1 =$ _____

The estimated standard error of b_1 is $s\{b_1\} =$ _____

The observed value of the test statistic for the t test of

$$H_0 : \beta_1 = 0$$

$$H_a : \beta_1 \neq 0$$

is $t =$ _____

The p-value = _____

Is the observed b_1 statistically significantly different from zero (Yes/No)? _____

What does this indicate about whether there's any relationship between the **murder rate** and the **illiteracy rate**?

Give the value of the *coefficient of determination* R^2 (labeled Multiple R-squared in the output from `summary()`).

Coefficient of determination $R^2 =$ _____

Based on the value of R^2 , what percentage of the variation in **murder rates** is explained by **illiteracy rate**? _____

6. *Don't* print the histogram, just answer the following question. Based on the histogram, does the assumption of normally distributed errors ϵ_i appear to be met? (Yes/No)?

7. *Don't* print the plot, just answer the following question. Based on the plot of residuals versus fitted values, does the assumption of a **constant standard deviation assumption** for ϵ appear to be met (Yes/No)? _____

8.

a) The **regression ANOVA table** in R gives the following:

Analysis of Variance Table						
Response: Y						
	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
X	Df for SSR	SSR	MSR	F	p	
Residuals	Df for SSE	SSE	MSE			

From **regression ANOVA table**, give the values of the following statistics:

SSR = _____ SSE = _____ SSTO = _____

MSR = _____ MSE = _____ F Test Statistic = _____

b) Give the degrees of freedom for SSR: _____

Give the degrees of freedom for SSE: _____

Give the degrees of freedom for SSTO: _____

c) Show that the F test statistic is the square of the t statistic from Step 5, and confirm that the p-values for the two tests are the same.

d) Show that

$$R^2 = 1 - \frac{\text{SSE}}{\text{SSTO}}$$

where R^2 is the coefficient of determination from Step 5.

2 Part B: Intercept-Only Regression

2.1 Murder Rates Data Set (Continued)

1. NA
2. Give the following values:

The estimate of β_0 is $b_0 =$ _____

The estimated standard error of b_0 is $s\{b_0\} =$ _____

Give the value of the *coefficient of determination* R^2 (labeled Multiple R-squared in the output from `summary()`).

Coefficient of determination $R^2 =$ _____

For which model, the one that included **illiteracy** as a predictor or the **intercept-only** model, is the R^2 larger?

3. *Don't* print the plot, just describe the fitted line.
4. Compare the value of \bar{Y} to that of b_0 . Are they the same (Yes/No)? _____
5. Are the results (*t* value and **p-value**) of the test for μ the same as those of the test for β_0 (Yes/No)? _____

3 Part C: Simple Linear Regression

3.1 Murder Rates Data Set (Continued)

1. NA
2. Give the endpoints of the *95% confidence interval for β_1* , and interpret the interval.
3. Give the value of the *estimate* of the mean response.
4. Give the endpoints of the *95% confidence interval for $E(Y_h)$* and interpret the interval.
5. Give the *95% prediction interval* for the **new response** $Y_{h(new)}$, and interpret the prediction interval.