

Homework Assignment Week 5

Tengyuan Liang
Business Statistics
Booth School of Business

Problem 1

A company sets different prices for a particular stereo system in eight different regions of the country. The table below shows the numbers of units sold (in 1000s of units) and the corresponding prices (in hundreds of dollars).

Sales	420	380	350	400	440	380	450	420
Price	5.5	6.0	6.5	6.0	5.0	6.5	4.5	5.0

- (i) In Excel, regress sales on price and obtain the intercept and slope estimates.

SUMMARY OUTPUT

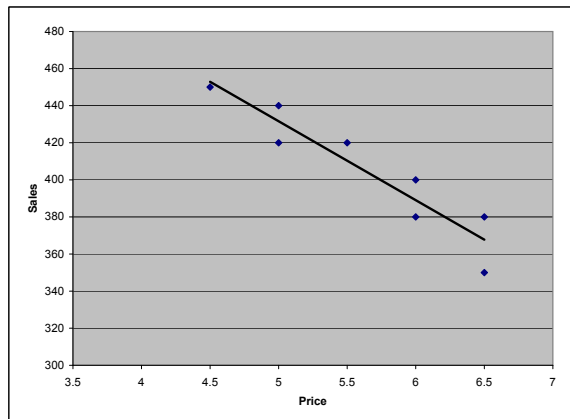
<i>Regression Statistics</i>	
Multiple R	0.937137027
R Square	0.878225806
Adjusted R Square	0.857930108
Standard Error	12.74227575
Observations	8

ANOVA

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	7025.806452	7025.806452	43.27152318	0.000592135
Residual	6	974.1935484	162.3655914		
Total	7	8000			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	644.516129	36.68873299	17.56714055	2.18343E-06	554.7420336	734.2902244
X Variable 1	-42.58064516	6.473082556	-6.578109392	0.000592135	-58.41970755	-26.74158277

- (ii) Present a plot with the data and the regression line



(iii) Based on this analysis, briefly describe your understanding of the relationship between sales and prices.

Problem 2: Match the Plots

Below (Figure 1) are 4 different scatter plots of an outcome variable y versus predictor x followed by 4 regression output summaries labeled A, B, C and D. Match the outputs with the plots.

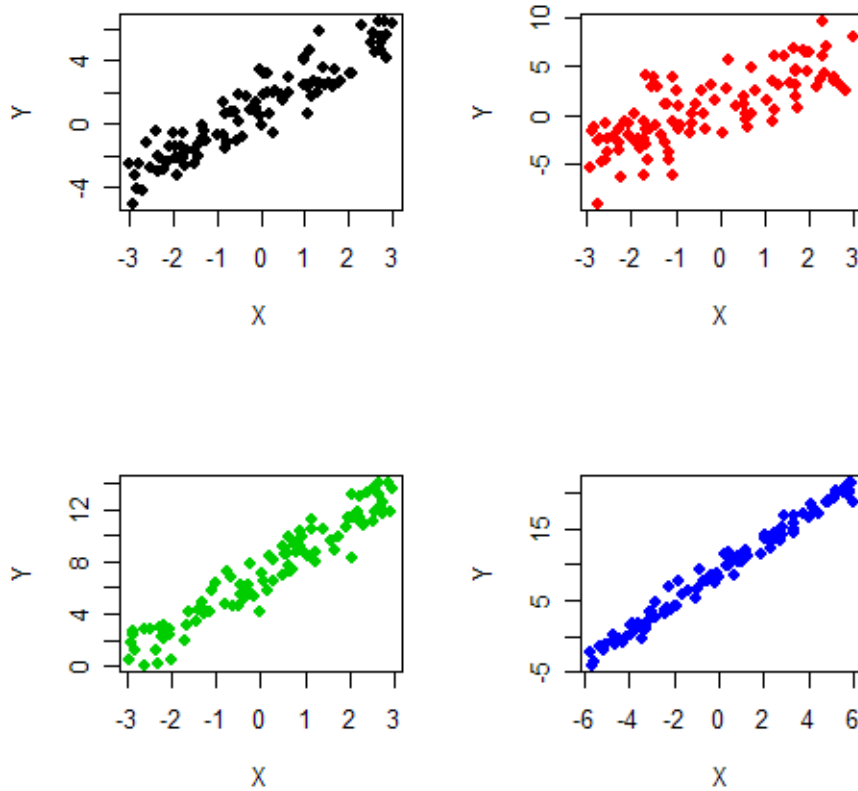


Figure 1: Scatter Plots

B-Black, C-Red, A-Green, D-Blue. Note: you can look at the R^2 , intercept and slope to decide how the outputs and plots match each other.

Regression A:

Coefficients:

	Estimate	Std. Error
(Intercept)	7.03747	0.12302
(Slope)	2.18658	0.07801

Residual standard error: 1.226

R-Squared: 0.8891

Regression B:

Coefficients:

	Estimate	Std. Error
(Intercept)	1.1491	0.1013
(Slope)	1.4896	0.0583

Residual standard error: 1.012

R-Squared: 0.8695

Regression C:

Coefficients:

	Estimate	Std. Error
(Intercept)	1.2486	0.2053
(Slope)	1.5659	0.1119

Residual standard error: 2.052

R-Squared: 0.6666

Regression D:

Coefficients:

	Estimate	Std. Error
(Intercept)	9.0225	0.0904
(Slope)	2.0718	0.0270

Residual standard error: 0.902

R-Squared: 0.9835

Problem 3

Suppose we are modeling house price as depending on house size. Price is measured in thousands of dollars and size is measured in thousands of square feet.

Suppose our model is:

$$P = 20 + 50s + \epsilon, \quad \epsilon \sim N(0, 15^2).$$

- (a) Given you know that a house has size $s = 1.6$, give a 95% predictive interval for the price of the house.

The point prediction is $\hat{P}_f = 20 + 50 \times 1.6 = 100$

The prediction interval is $[100 \pm 2 \times 15] = [70; 130]$

- (b) Given you know that a house has size $s = 2.2$, give a 95% predictive interval for the price.

The point prediction is $\hat{P}_f = 20 + 50 \times 2.2 = 130$

The prediction interval is $[130 \pm 2 \times 15] = [100; 160]$

- (c) In our model the slope is 50. What are the units of this number?

1,000\$ / 1,000 Sq. Feet = \$/Sq. Feet

- (d) What are the units of the intercept 20?

1,000\$ (same as P)

- (e) What are the units of the the error standard deviation 15?

1,000\$ (same as P)

- (f) Suppose we change the units of price to dollars and size to square feet
What would the values and units of the intercept, slope, and error standard deviation?

Intercept: 20,000 \$

Slope: 50 \$/Sq. Feet

error standard deviation: 15,000 \$

- (g) If we plug $s = 1.6$ into our model equation, P is a constant plus the normal random variables ϵ . Given $s = 1.6$, what is the distribution of P ?

When $s = 1.6$ the mean of house prices is $20 + 50 \times 1.6 = 100$. The error standard deviation is the same, 15. Therefore

$$P|s = 1.6 \sim N(100, 15^2)$$