WEEK 7 HOMEWORK

1. A real estate analyst has developed a multiple regression line, $y = 60 + 0.068\ x_1 - 2.5\ x_2$, to predict y = the market price of a home (in $1,000s), using independent variables, x_1 = the total number of square feet of living space, and x_2 = the age of the house in years. The regression coefficient of x_2 suggests this: _____.

- Whatever be the square feet area of the living space, a 1 year increase in the age of the homes will result in a predicted drop of $2500 in the price of the homes
- If the square feet area of living space is kept constant, a 1 year increase in the age of the homes will result in a predicted increase of $2500 in the price of the homes
- If the square feet area of living space is kept constant, a 1 year increase in the age of the homes will result in a predicted drop of $2500 in the price of the homes
- Whatever be the square feet area of the living space, a 1 year increase in the age of the homes will result in a predicted increase of $2500 in the price of the homes

2. The following is a **partial** computer output of a multiple regression analysis of a data set containing 20 sets of observations on the dependent variable

The regression equation is
SALEPRIC = 1470 + 0.8145 LANDVAL + 0.8204 IMPROVAL + 13.529 AREA

Predictor	Coef	SE Coef	T	P
Constant	1470	5746	0.26	0.801
LANDVAL	0.8145	0.5122	1.59	0.131
IMPROVAL	0.8204	0.2112	3.88	0.0001
AREA	13.529	6.586	2.05	0.057

S = 79190.48 R-Sq = 89.7% R-Sq(adj) = 87.8%

Analysis of Variance

Source	DF	SS	MS
Regression	3	8779676741	2926558914
Residual Error	16	1003491259	62718204
Total	19	9783168000	

What is the prediction value when LANDVAL = 25, IMPROVAL = 36 and

AREA = 56?

- 2277.521
- 2547.521
- 2500.25
- 2200.78

3. A cost accountant is developing a regression model to predict the total cost of producing a batch of printed circuit boards as a linear function of batch size (the number of boards produced in one lot or batch), production plant (Kingsland, and Yorktown), and production shift (day, and evening). The response variable in this model is ___.

- production plant
- production shift
- batch size
- variable cost
- total cost

4. A cost accountant is developing a regression model to predict the total cost of producing a batch of printed circuit boards as a linear function of batch size (the number of boards produced in one lot or batch), production plant (Kingsland, and Yorktown), and production shift (day, and evening). In this model, "batch size" is_____.

- a response variable
- a dependent variable
- a qualitative variable
- an indicator variable
- an independent variable

5. To predict the value of the dependent variable of our interest (*y*) for the given, specific values of the independent variables (the *x*-

variables) using a multiple regression model, we substitute the values of the *x*-variables into the multiple regression equation and solve for the value of *y*. The resulting value of *y* is_____.

- an exact forecast of *y*
- a confidence interval of the average value of *y*
- a point estimate of *y*
- a prediction interval of for a single value of *y*

6. Suppose that the regression equation $y = 16.99 + 0.32\ x1 + 0.41\ x2 + 5.31\ x3$ predicts an adult's height (*y*) given the individual's mother's height (*x*1), his or her father's height (*x*2), and whether the individual is male (*x*3 = 1) or female (*x*3 = 0). All heights are measured in inches. Assume also that this equation is stable through time, the average adult female height is currently 63.8 inches and the average adult male height is 69.7 inches. Approximately what would be the average female height in two generations? You can assume that each individual has parents of average height.

- 67.25 inches
- 67.33 inches
- 67.82 inches
- 66 inches.
- There is not enough information to determine the average female height in two generations.

7. A market analyst is developing a regression model to predict monthly household expenditures on groceries as a function of family size, household income, and household neighborhood (urban, suburban, and rural). The "neighborhood" variable in this model is _____.

- a quantitative variable
- a dependent variable

○ a response

variable ○ a constant

○ an independent variable

8. The following is a **partial** computer output of a multiple regression analysis of a data set containing 20 sets of observations on the dependent variable

The regression equation is
SALEPRIC = 1470 + 0.814 LANDVAL + 0.820 IMPROVAL + 13.5 AREA

Predictor	Coef	SE Coef	T	P
Constant	1470	5746	0.26	0.801
LANDVAL	0.8145	0.5122	1.59	0.131
IMPROVAL	0.8204	0.2112	3.88	0.0001
AREA	13.529	6.586	2.05	0.057

S = 79190.48 R-Sq = 89.7% R-Sq(adj) = 87.8%

Analysis of Variance

Source	DF	SS	MS
Regression	3	8779676741	2926558914
Residual Error	16	1003491259	62718204
Total	19	9783168000	

For the problem above, we want to carry out the significance test about the coefficient of LANDVAL, what is the t-value for this test, and is it significant?

○ 0.26, not significant
○ 46.66, significant
○ 1.59, not significant
○ 2.05, significant

9. In testing the overall significance of a multiple regression model, the null hypothesis that each one the ß-coefficients of the *x*-variables in the model is equal to zero, is tested against the alternative hypothesis that_____.

○ at least two of the ß-coefficients of the *x*-variables in the model is ≠ zero

○ at least one of the ß-coefficients of the *x*-variables in the model is ≠ zero

○ each one of the ß-coefficients of the *x*-variables in the model is > zero

○ each one of the ß-coefficients of the *x*-variables in the model is ≠ zero

10. In regression analysis, outliers may be identified by examining the _____.

○ *p*-values for the partial coefficients

○ *R*-squared value

○ coefficient of correlation

○ coefficient of determination

○ residuals