MGS 8040: Regression Analysis Exercise

**Consider the following output** of a regression model to predict ***home prices*** *($ thousand)* in a region based on *area of the house, age of the house, number of bedrooms* and *number of bathrooms*:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT | |  |  |  |  |  |
|  |  | |  |  |  |  |
| *Regression Statistics* | | |  |  |  |  |
|  |  | |  |  |  |  |
| R Square |  | |  |  |  |  |
|  |  | |  |  |  |  |
| Standard Error |  | |  |  |  |  |
| Observations | 20 | |  |  |  |  |
|  |  | |  |  |  |  |
| ANOVA |  | |  |  |  |  |
|  | *df* | | *SS* | *MS* | *F* | *Significance F* |
| Regression |  | | 322.3529 | 80.58822 |  | 0.002804 |
| Residual |  | |  | 12.16314 |  |  |
| Total |  | | 504.8000 |  |  |  |
|  |  | |  |  |  |  |
|  | *Coefficients* | | *Standard Error* | *t Stat* | *P-value* |  |
| Intercept | 215.000 | | 3.57732 | 4.314106 | 0.000614 |  |
| Age | -1.000 | | 0.372639 | -2.80578 | 0.033304 |  |
| Area | 0.050 | | 0.023784 | 2.958883 | 0.009756 |  |
| Bedrooms | 9.000 | | 2.084543 | 3.082656 | 0.007581 |  |
| Bathrooms | 13.000 | | 1.000000 | 3.000000 | 0.022561 |  |

**Fill the shaded cells above.**

1. The **R-Square value** for the regression is equal to: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

It means that \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. The **Standard Error** is equal to \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
2. Is the Regression significant at the 5% level**? \_\_\_\_\_\_\_\_\_\_\_\_\_**At 1%? **\_\_\_\_\_\_\_\_\_\_\_\_\_\_**
3. **The F value** for the regression is \_\_\_\_\_\_\_\_\_\_\_\_\_\_
4. The **coefficient -1.00 for age** means that \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. Based on the regression above, the **predicted value** (in thousands of dollars) for a new home (age=0) that is 2000 sq. ft in area, with 3 bedrooms and 3 bathrooms is

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. The **Margin of Error** for the 95% prediction interval for your prediction above is approximately

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. You are testing the relationship between some X and Y for three different materials, A, B, and C. You wish to account for the material used in your regression analysis. Define the dummy variable values for a few observations of material type shown below:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Y** | **X** | **Material** | **Dummy A** | **Dummy B** |
| 20 | 7 | A |  |  |
| 18 | 6 | B |  |  |
| 15 | 5 | B |  |  |
| 14 | 4 | C |  |  |
| 18 | 5 | C |  |  |
| 17 | 6 | A |  |  |
| 11 | 5 | A |  |  |