12.32 *t*-score? A regression analysis is conducted with 25 observations.

a. What is the *df* value for inference about the slope β?

b. Which two *t* test statistic values would give a P-value of 0.05 for testing H_0: $\beta = 0$ against H_a: $\beta \neq 0$?

c. Which *t*-score would you multiply the standard error by in order to find the margin of error for a 95% confidence interval for β?

a. n=25, df=n-2=25-2=23.

b. t= ± 2.069

c. 2.069

12.33 Predicting house prices For the House Selling Prices FL data file on the text CD, MINITAB results of a regression analysis are shown for 100 homes relating y = selling price (in dollars) to x = the size of the house (in square feet).

a. Show all steps of a two-sided significance test of independence. Could the sample association between these two variables be explained by random variation?

b. Show that a 95% confidence interval for the population slope is (64, 90).

c. A builder had claimed that the selling price increases $100, on average, for every extra square foot. Based on part b, what would you conclude about this claim?

House selling prices and size of home

Predictor	Coef	SE Coef	T	P
Constant	9161	10760	0.85	0.397
size	77.008	6.626	11.62	0.000

12.34 House prices in bad part of town Refer to the previous exercise. Of the 100 homes, 25 were in a part of town considered less desirable. For a regression analysis using y = selling price and x = size of house for these 25 homes,

a. You plan to test H_0: $\beta = 0$ against H_a: $\beta > 0$. Explain what H_0 means, and explain why a data analyst might choose a one-sided H_a for this test.

b. For this one-sided alternative hypothesis, how large would the *t* test statistic need to be in order to get a P-value equal to (i) 0.05 and (ii) 0.01?

a. There is no association between selling price and size of house.

A right-tailed test can be used since previous analysis showed positive association between x and y.

b. n=25, df=23. (i) t=1.714 (ii) t=2.500.

a. 1.Linear trend, random, normal conditional distribution for y at each x value with same std. dev..

2. $H_0: \beta = 0 : H_a: \beta \neq 0$

3. t=(b-0)/se(b)=77.008/6.626=11.622

4. P-value=2P(t>11.622)=0.000

5. If H_0 were true, there is almost 0 probability of getting a result at least as extreme as the value observed. Since the P-value is extremely small, we reject H_0 and conclude that an positive association exists between the size and price of houses.

b. 95% CI is b ± 1.987 se(b)=(64,90)

c. $\beta = 100$ is not in the interval. It is not plausible to have slope.

12.42 Student GPAs Refer to the Georgia Student Survey data file on the text CD. Treat college GPA as the response variable and high school GPA as the explanatory variable, and suppose these students are a random sample of all University of Georgia students. **n=59**

a. Can you conclude that these variables are associated in that population? Show all steps of the relevant significance test with significance level 0.05, and interpret.

b. Find a 95% confidence interval for the population slope. Interpret the endpoints, and explain the correspondence with the result of the significance test in part a.

```
The regression equation is
CGPA = 1.19 + 0.637 HSGPA

Predictor     Coef   SE Coef      T       P
Constant    1.1898    0.5496   2.16   0.035
HSGPA       0.6369    0.1442   4.42   0.000
```

a. 1.Linear trend, random, normal conditional distribution for y at each x value with same std. dev..

2. $H_0: \beta = 0 : H_a: \beta \neq 0$

3. t=(b-0)/se(b)=0.6369/0.1442=4.42

4. P-value=2P(t>4.42)=0.000

5. If H_0 were true, there is almost 0 probability of getting a result at least as extreme as the value observed. Since the P-value is extremely small, we reject H_0 and conclude that a positive association exists between the HSGPA and CGPA.

b. n=59, df=57. 95% CI is b $\pm$ 2.009 se(b)=(0.3, 0.9)

0 is not in the interval. All plausible values are positive. As we concluded in the significance test, there is a positive association between HSGPA and CGPA.

12.54 Predicting GPA Refer to the Georgia Student Survey data file on the text CD. Regress y = college GPA on x = high school GPA.

a. Stating the necessary assumptions, find a 95% confidence interval for the mean college GPA for all University of Georgia students who have high school GPA = 3.6.

b. Find a 95% prediction interval for college GPA of a randomly chosen student having high school GPA = 3.6. Interpret.

c. Explain the difference between the purposes of the intervals in part a and part b.

```
Predicted Values for New Observations
New
Obs   Fit    SE Fit        95% CI              95% PI
1    3.4828 0.0505  (3.3818, 3.5839) (2.8407, 4.1250)

Values of Predictors for New Observations
New
Obs   HSGPA
  1    3.60
```

a. Linear trend, random, normal conditional distribution for y at each x value with same std. dev..

95% CI for μ_y when x=3.6 is $(3.38, 3.58)$

b. 95% PI for y when x=3.6 is $(2.84, 4.13)$

c. A 95% PI is used to predict individual CGPA to fall for a given value of x.

A 95% CI is used to give plausible values of the population mean of CGPA at a given x.

PI is wider than CI.