Student name:\_\_\_\_\_\_\_\_\_\_

**1)** The following output resulted from a regression model where SAGap is seasonally adjusted Gap sales and dpi is disposable income per capita.

|  |
| --- |
| Audit Trail -- Coefficient Table (Mulitple Regression Selected) |
| Series Description | Included in Model  | Coefficient  | Standard Error  | T-test  | P-value  | F-test  | Elasticity  |
| SAGAP |  | Dependent |  | - | 2,867,564.78 |  |  | 140,536.33 |  | - | 20.40 |  |  | 0.00 |  |  | 416.34 |  |  |  |  |
| dpi |  | Yes |  |  | 809.79 |  |  | 25.04 |  |  | 32.33 |  |  | 0.00 |  |  | 1,045.55 |  |  | 2.91 |  |
|  |

|  |
| --- |
| Audit Trail -- Correlation Coefficient Table |
| Series Description | Included in Model | SAGap | dpi |
| SAGap |  | Dependent |  |  | 1.00 |  |  | 0.97 |  |
| dpi |  | Yes |  |  | 0.97 |  |  | 1.00 |  |
|  |

|  |  |
| --- | --- |
| Audit Trail - Statistics |  |
| Accuracy Measures | Value |  | Forecast Statistics | Value |  |
| AIC | 2,135.23 |  | Durbin Watson(4) | 0.43 |  |
| BIC | 2,137.56 |  | Mean | 1,501,041.15 |  |
| Mean Absolute Percentage Error (MAPE) | 34.60 | % | Standard Deviation | 1,303,264.45 |  |
| R-Square | 94.59 | % | Max | 4,253,174.95 |  |
| Adjusted R-Square | 94.51 | % | Min | 123,121.77 |  |
| Mean Square Error | 90,711,613,878.48 |  | Range | 4,130,053.18 |  |
| Root Mean Square Error | 301,183.69 |  | Root Mean Square | 1,294,661.95 |  |
| Theil | 6.23 |  | Ljng-Box | 347.40 |  |
|  |  |

 A) None of the options are true.
 B) This regression model is a lagged model.
 C) This regression model is a nonlinear model.
 D) This regression model is a causal model.
 E) This regression model is a multiple regression model.

**2)** The following regression represents seasonally adjusted shoe store sales in millions of dollars in the United States.

|  |
| --- |
| Audit Trail -- ANOVA Table (Multiple Regression Selected) |
| Source of variation | SS | df | MS | SEE |
| Regression |  | 2,099,394.07 |  |  | 1 |  |  | 2,099,394.07 |  |  |  |  |
| Error |  | 632,943.91 |  |  | 142 |  |  | 4,457.35 |  |  | 66.76 |  |
| Total |  | 2,732,337.98 |  |  | 143 |  |  |  |  |  |  |  |
|  |

|  |
| --- |
| Audit Trail -- Coefficient Table (Mulitple Regression Selected) |
| Series Description | Included in Model  | Coefficient  | Standard Error  | T-test  | P-value  | F-test  | Elasticity  | Overall F-test  |
| SASSS |  | Dependent |  |  | 1,523.74 |  |  | 11.19 |  |  | 136.23 |  |  | 0.00 |  |  | 18,557.38 |  |  |  |  |  | 471.00 |  |
| Time index |  | Yes |  |  | 2.90 |  |  | 0.13 |  |  | 21.70 |  |  | 0.00 |  |  | 471.00 |  |  | 0.12 |  |  |  |  |
|  |

|  |
| --- |
| Audit Trail - Statistics |
| Accuracy Measures | Value |  | Forecast Statistics | Value |
| AIC | 1,618.57 |  | Durbin Watson(12) | 1.15 |
| BIC | 1,621.54 |  | Mean | 1,734.33 |
| Mean Absolute Percentage Error (MAPE) | 3.06 | % | Standard Deviation | 138.23 |
| R-Square | 76.84 | % | Max | 1,988.06 |
| Mean Absolute Error | 51.97 |  | Min | 1,397.82 |
| Mean Square Error | 4,395.44 |  | Mean Absolute Deviation | 114.40 |
| Root Mean Square Error | 66.30 |  | Mean Square Deviation | 18,974.57 |
| Theil | 0.80 |  | Ljng-Box | 85.17 |
|  |

|  |  |
| --- | --- |
| Method Statistics | Value |
| Method Selected | Multiple Regression |
|  |

Consider the simple regression above. The dependent variable is seasonally adjusted shoe store sales in the United States. The independent variable is the index of time.The slope term in this regression

 A) is 1523.74.
 B) is 11.19.
 C) is 2.90.
 D) is 0.13.

**3)** The Y-intercept of the simple regression model

 A) is always positive when the correlation between the dependent and independent variable is positive.
 B) rarely has a useful interpretation.
 C) almost always has a useful interpretation.
 D) All of the options are correct.
 E) is always a positive number.

**4)** The following regression represents seasonally adjusted shoe store sales in millions of dollars in the United States.

|  |
| --- |
| Audit Trail -- ANOVA Table (Multiple Regression Selected) |
| Source of variation | SS | df | MS | SEE |
| Regression |  | 2,099,394.07 |  |  | 1 |  |  | 2,099,394.07 |  |  |  |  |
| Error |  | 632,943.91 |  |  | 142 |  |  | 4,457.35 |  |  | 66.76 |  |
| Total |  | 2,732,337.98 |  |  | 143 |  |  |  |  |  |  |  |
|  |

|  |
| --- |
| Audit Trail -- Coefficient Table (Mulitple Regression Selected) |
| Series Description | Included in Model  | Coefficient  | Standard Error  | T-test  | P-value  | F-test  | Elasticity  | Overall F-test  |
| SASSS |  | Dependent |  |  | 1,523.74 |  |  | 11.19 |  |  | 136.23 |  |  | 0.00 |  |  | 18,557.38 |  |  |  |  |  | 471.00 |  |
| Time index |  | Yes |  |  | 2.90 |  |  | 0.13 |  |  | 21.70 |  |  | 0.00 |  |  | 471.00 |  |  | 0.12 |  |  |  |  |
|  |

|  |
| --- |
| Audit Trail - Statistics |
| Accuracy Measures | Value |  | Forecast Statistics | Value |
| AIC | 1,618.57 |  | Durbin Watson(12) | 1.15 |
| BIC | 1,621.54 |  | Mean | 1,734.33 |
| Mean Absolute Percentage Error (MAPE) | 3.06 | % | Standard Deviation | 138.23 |
| R-Square | 76.84 | % | Max | 1,988.06 |
| Mean Absolute Error | 51.97 |  | Min | 1,397.82 |
| Mean Square Error | 4,395.44 |  | Mean Absolute Deviation | 114.40 |
| Root Mean Square Error | 66.30 |  | Mean Square Deviation | 18,974.57 |
| Theil | 0.80 |  | Ljng-Box | 85.17 |
|  |

|  |  |
| --- | --- |
| Method Statistics | Value |
| Method Selected | Multiple Regression |
|  |

Consider the simple regression above. The dependent variable is seasonally adjusted shoe store sales in the United States. The independent variable is the index of time.The reported p value for the time index

 A) indicates that there is little explanatory power to this regression model.
 B) indicates that the time variable is statistically significant.
 C) indicates that the time variable is unrelated to shoe sales.
 D) indicates that the time variable is statistically insignificant.

**5)** In table 4.1 from the text, four different data sets are displayed along with the regressions associated with each data set. What point was being made in the text?

 A) While each data set is quite different from the rest, all result in the same regression.
 B) Regression works equally well with almost any data set.
 C) While there are four different data sets, all differ only by some random variation.
 D) All the data sets are more alike than they are different.

**6)** There is an approximate linear relationship between the height of females and their age (from 5 to 18 years) described by:height = 50.3 + 6.01(age)where height is measured in cm and age in years. Which of the following is not correct?

 A) The estimated slope is 6.01 which implies that children increase by about 6 cm for each year they grow older.
 B) My niece is about 8 years old and is about 115 cm tall. She is taller than average.
 C) The average height of children when they are 5 years old is about 50% of the average height when they are 18 years old.
 D) The estimated intercept is 50.3 cm which implies that children reach this height when they are 50.3/6.01=8.4 years old.
 E) The estimated height of a child who is 10 years old is about 110 cm.

**7)** The Durbin-Watson statistic

 A) tests for serial probability.
 B) tests the coefficient(s) of the independent variable(s).
 C) takes on values from 0 to 4.
 D) cannot be used with fewer than 10 observations.

**8)** What pattern of residuals in a time-series plot would indicate heteroscedasticity?

 A) Circular
 B) Purely random
 C) Funnel shaped
 D) Linear
 E) Elliptical

**9)**

|  |
| --- |
| Audit Trail -- ANOVA Table (Trend (Linear) Regression Selected) |
| Source of variation | SS | df | MS | SEE |
| Regression |  | 149,557,032.66 |  |  | 1 |  |  | 149,557,032.66 |  |  |  |  |
| Error |  | 1,006,968.39 |  |  | 130 |  |  | 7,745.91 |  |  | 88.01 |  |
| Total |  | 150,564,001.05 |  |  | 131 |  |  |  |  |  |  |  |
|  |

|  |
| --- |
| Audit Trail — Coefficient Table (Trend (Linear) Regression Selected) |
| Name | Value  | Standard Error | T-test | Elasticity  | Overall F-test |
| Intercept |  | 4,588.58 |  |  | 15.41 |  |  | 297.80 |  |  |  |  |  | 19,307.87 |  |
| Slope |  | 27.93 |  |  | 0.20 |  |  | 138.95 |  |  | 0.29 |  |  |  |  |
|  |

|  |
| --- |
| Audit Trail - Statistics |
| Accuracy Measures | Value | Forecast Statistics | Value |
| AIC | 1,558.63 |  | Durbin Watson(1) | 0.21 |  |
| BIC | 1,564.40 |  | Mean | 6,446.25 |  |
| Mean Absolute Percentage Error (MAPE) | 1.24 | % | Standard Deviation | 1,072.07 |  |
| R-Square | 99.33 | % | Max | 8,364.50 |  |
| Adjusted R-Square | 99.33 |  | Min | 4,800.10 |  |
| Root Mean Square Error | 87.34 |  | Range | 3,564.40 |  |
| Theil | 1.89 |  | Root Mean Square | 1,068.01 |  |
|  |  |  | Ljung-Box | 543.24 |  |
|  |

|  |  |
| --- | --- |
| Method Statistics | Value |
| Method Selected | Trend (Linear) Regression |
|  |

This is a linear trend regression. Which of the following is not true?

 A) There is a near vertical linear trend.
 B) No improvement is shown from that of a naïve model since the Theil’s U is just 1.89.
 C) There is some explanatory power since the MAPE is only 1.24%.
 D) There is a negatively sloped linear trend.

**10)** Confidence Bands

Given that we have collected pairs of observations on two variables X and Y, we would consider fitting a straight line with X as an explanatory variable if:

 A) the change in Y is a constant for each unit change in X.
 B) the change in Y is a fixed percent of Y.
 C) the change in Y is exponential.
 D) None of the options are correct.
 E) the change in Y is an additive constant.

**11)** If the residuals in a regression equation are positively autocorrelated, which of the following is not a problem when the least squares procedure is used?

 A) The regression coefficients are no longer strictly applicable.
 B) Confidence intervals are no longer strictly applicable.
 C) The standard error of the regression seriously understates the variability of the error terms.
 D) The t and F distributions are no longer strictly applicable.
 E) The standard error of the regression slope coefficient underestimates the true variability of the estimated regression.

**12)** If R2 is .95 in a simple regression model, it can be said that

 A) 5 percent of Y's variability is caused by variability in X.
 B) X and Y have a correlation of .95.
 C) 95% of Y's variability can be explained by X's variability.
 D) the relationship between X and Y is positive.

**13)** The standard error of the estimate in an ordinary least squares regression is a measure of the

 A) explained variation.
 B) total variation of the Y variable.
 C) variation of the X variable.
 D) variation around the sample regression line.

**14)** In a cross-sectional study of sales in different cities, the following relationship between sales revenue (S = sales revenue in dollars) and city size as measured by population (POP = population in thousands) was estimated: S = 37.02 + 0.6734(POP). What is the approximate 95 percent prediction interval for sales in a city of 155,000 people when the regression standard error is 32.7214? (Note: you should use a rule of thumb here)

 A) 46.721 to 254.586.
 B) 95.899 to 187.333.
 C) 75.955 to 206.839.
 D) 108.676 to 174.118.

**15)** The Standard Error of the Estimate (SEE) in an ordinary least squares regression

 A) is used to test for autocorrelation.
 B) is the standard error of the independent variable.
 C) is used to test for serial correlation.
 D) is the standard error of the dependent variable.

**16)** Which of the following is incorrect?

 A) R2 is a measure of the degree of variability in the dependent variable about its sample mean explained by the regression line.
 B) None of the options are correct.
 C) R2 measures the "goodness-of-fit" of a regression model.
 D) The null hypothesis that R2 = 0 can be tested using an F-test.
 E) The best model selection criteria and variable selection criteria for a forecaster to use is the maximization of R Squared.

**17)** A study was conducted to examine the quality of fish after seven days in ice storage. For this studyY = measurement of fish quality (on a 10-point scale with 10 = BEST.)X = # of hours after being caught that the fish were packed in ice.The sample linear regression line is: Y = 8.5 - .5X. From this we can say that:

 A) Can’t really say until we see a plot of the data.
 B) a one hour delay in packing the fish in ice decreases the estimated quality by .5.
 C) if the estimated quality increases by 1, then the fish have been packed in ice one hour sooner.
 D) if the estimated quality increases by 1, then the fish have been packed in ice two hours later.
 E) a one hour delay in packing the fish in ice increases the estimated quality by .5.

**18)** Exact prediction intervals for the dependent variable

 A) are linear around the estimated regression line.
 B) are bow-shaped around the estimated regression line.
 C) None of the options are correct.
 D) do not take the variability of Y around the sample regression into account.
 E) do not take the randomness of the sample into account.

**19)** Which of the following would indicate a perfect model fit?

 A) R2 = 0
 B) t-test for slope > 2
 C) Durbin-Watson = 2
 D) None of the options are correct.
 E) R2 = 1

**20)** The Durbin-Watson statistic is based upon the

 A) All of the options are correct.
 B) estimated residuals.
 C) estimated slope coefficient.
 D) estimated Y-intercept term.
 E) fitted regression line.

**21)** What is the approximate 95% prediction interval for the dependent variable when the independent variable value is 20, assuming the fitted regression line is: Y = 1.50 + 6.0(X). Assume the sample size is 20 and the standard error of the regression (SEE) is 1.2. You should use the "rule of thumb" used in class here.

 A) ~17.65 to 22.35
 B) ~119.32 to 124.12
 C) ~117.84 to 126.16
 D) ~108.60 to 137.40
 E) ~119.10 to 123.90

**22)** The coefficient of determination in an ordinary least squares regression tells us

 A) whether the correlation coefficient has any significance.
 B) that the correlation coefficient is larger than 1.
 C) the proportion of the total variation that is explained.
 D) that we should not partition the total variation.

**23)** Testing the null hypothesis that the slope coefficient is zero uses what sampling distribution for small sample sizes?

 A) Standard Normal.
 B) None of the options are correct.
 C) t distribution with n-1 degrees of freedom.
 D) Normal.
 E) Chi-square.

**24)** Sample regression model forecast errors are called

 A) None of the options are correct.
 B) least-squares predictions.
 C) outliers.
 D) residuals.
 E) disturbances.

**25)** The following regression represents seasonally adjusted shoe store sales in millions of dollars in the United States.

|  |
| --- |
| Audit Trail -- ANOVA Table (Multiple Regression Selected) |
| Source of variation | SS | df | MS | SEE |
| Regression |  | 2,099,394.07 |  |  | 1 |  |  | 2,099,394.07 |  |  |  |  |
| Error |  | 632,943.91 |  |  | 142 |  |  | 4,457.35 |  |  | 66.76 |  |
| Total |  | 2,732,337.98 |  |  | 143 |  |  |  |  |  |  |  |
|  |

|  |
| --- |
| Audit Trail -- Coefficient Table (Mulitple Regression Selected) |
| Series Description | Included in Model  | Coefficient  | Standard Error  | T-test  | P-value  | F-test  | Elasticity  | Overall F-test  |
| SASSS |  | Dependent |  |  | 1,523.74 |  |  | 11.19 |  |  | 136.23 |  |  | 0.00 |  |  | 18,557.38 |  |  |  |  |  | 471.00 |  |
| Time index |  | Yes |  |  | 2.90 |  |  | 0.13 |  |  | 21.70 |  |  | 0.00 |  |  | 471.00 |  |  | 0.12 |  |  |  |  |
|  |

|  |
| --- |
| Audit Trail - Statistics |
| Accuracy Measures | Value |  | Forecast Statistics | Value |
| AIC | 1,618.57 |  | Durbin Watson(12) | 1.15 |
| BIC | 1,621.54 |  | Mean | 1,734.33 |
| Mean Absolute Percentage Error (MAPE) | 3.06 | % | Standard Deviation | 138.23 |
| R-Square | 76.84 | % | Max | 1,988.06 |
| Mean Absolute Error | 51.97 |  | Min | 1,397.82 |
| Mean Square Error | 4,395.44 |  | Mean Absolute Deviation | 114.40 |
| Root Mean Square Error | 66.30 |  | Mean Square Deviation | 18,974.57 |
| Theil | 0.80 |  | Ljng-Box | 85.17 |
|  |

|  |  |
| --- | --- |
| Method Statistics | Value |
| Method Selected | Multiple Regression |
|  |

Consider the simple regression above. The dependent variable is seasonally adjusted shoe store sales in the United States. The independent variable is the index of time.Does it appear that there is a trend to shoe store sales?

 A) None of the options are true.
 B) Yes, the slope term is positive and statistically significant.
 C) No, although the slope term is positive, it is statistically insignificant.
 D) There is no way to tell from the data given whether there is a trend to shoe store sales.

**26)** Serial correlation (or autocorrelation) causes estimates of the

 A) estimates of the standard errors to be understated on average.
 B) None of the options are correct.
 C) slope parameter β to be overstated on average.
 D) slope parameter β to be understated on average.
 E) estimates of the standard errors to be overstated on average.

**27)** The autocorrelation parameter defined asis used to measure

 A) correlation between regression error terms.
 B) the Durbin-Watson statistic.
 C) the difference between the forecast and the estimated regression line.
 D) disturbances of independent random variables.
 E) None of the options are correct.

**28)** The following regression represents seasonally adjusted shoe store sales in millions of dollars in the United States.

|  |
| --- |
| Audit Trail -- ANOVA Table (Multiple Regression Selected) |
| Source of variation | SS | df | MS | SEE |
| Regression |  | 2,099,394.07 |  |  | 1 |  |  | 2,099,394.07 |  |  |  |  |
| Error |  | 632,943.91 |  |  | 142 |  |  | 4,457.35 |  |  | 66.76 |  |
| Total |  | 2,732,337.98 |  |  | 143 |  |  |  |  |  |  |  |
|  |

|  |
| --- |
| Audit Trail -- Coefficient Table (Mulitple Regression Selected) |
| Series Description | Included in Model  | Coefficient  | Standard Error  | T-test  | P-value  | F-test  | Elasticity  | Overall F-test  |
| SASSS |  | Dependent |  |  | 1,523.74 |  |  | 11.19 |  |  | 136.23 |  |  | 0.00 |  |  | 18,557.38 |  |  |  |  |  | 471.00 |  |
| Time index |  | Yes |  |  | 2.90 |  |  | 0.13 |  |  | 21.70 |  |  | 0.00 |  |  | 471.00 |  |  | 0.12 |  |  |  |  |
|  |

|  |
| --- |
| Audit Trail - Statistics |
| Accuracy Measures | Value |  | Forecast Statistics | Value |
| AIC | 1,618.57 |  | Durbin Watson(12) | 1.15 |
| BIC | 1,621.54 |  | Mean | 1,734.33 |
| Mean Absolute Percentage Error (MAPE) | 3.06 | % | Standard Deviation | 138.23 |
| R-Square | 76.84 | % | Max | 1,988.06 |
| Mean Absolute Error | 51.97 |  | Min | 1,397.82 |
| Mean Square Error | 4,395.44 |  | Mean Absolute Deviation | 114.40 |
| Root Mean Square Error | 66.30 |  | Mean Square Deviation | 18,974.57 |
| Theil | 0.80 |  | Ljng-Box | 85.17 |
|  |

|  |  |
| --- | --- |
| Method Statistics | Value |
| Method Selected | Multiple Regression |
|  |

Consider the simple regression above. The dependent variable is seasonally adjusted shoe store sales in the United States. The independent variable is the index of timeThe coefficient of determination in the shoe store sales regression

 A) indicates that about 77 percent of the variation in shoe sales is explained.
 B) indicates that the independent variable is statistically significant.
 C) indicates that the confidence interval for this regression would be plus and minus two times 0.12.
 D) indicates that there is no serial correlation in the error terms.

**29)** In linear least squares regression, we use "squares" to overcome what difficulty?

 A) using squares reduces the size of the error terms in a regression analysis.
 B) using squares counts both positive and negative errors rather than allowing them to offset one another.
 C) because we often use time series data, using squares is required.
 D) because standard deviations will not be "standardized" unless squares are used.

**30)** What is not likely to be a problem when applying ordinary least squares to cross-sectional data?

 A) Nonlinearity.
 B) Unbiased coefficient estimates.
 C) Heteroscedasticity.
 D) None of the options are correct.
 E) Autocorrelation.

**31)** The following regression equation was estimated: Y = −2.0 + 4.6X. This indicates that

 A) All of the options are correct.
 B) Y equals 44 when X is 10.
 C) there has been an error since "b" cannot be a negative number.
 D) the correlation coefficient for Y and X will be negative.
 E) there is a negative relationship between the two variables.

**32)** Because every software package that includes regression analysis uses the "Normal Equations,"

 A) every package will include "Normalized Coefficients."
 B) every software package will provide slightly different results.
 C) every software package will provide identical results.
 D) all of the diagnostic statistics produced by the various packages will vary.

**33)** The yield of a grain, Y (t/ha), appears to be linearly related to the amount of fertilizer applied, X (kg/ha). An experiment was conducted by applying different amounts of fertilizer (0 to 10 kg/ha) to plots of land and measuring the resulting yields. The following estimated regression line was obtained:yield = 4.85 + .05(fertilizer)Which of the following is not correct?

 A) If no fertilizer was used, the yield is estimated to be 4.85 t/ha (note that zero use of fertilizer is within the range of this data).
 B) If the current level of fertilizer is changed from 7.0 to 9.0 kg/ha, the yield is estimated to increase by 0.20 t/ha.
 C) For every additional kg/ha of fertilizer applied, the yield is estimated to increase 0.05 t/ha.
 D) If fertilizer is applied at 10 kg/ha, the estimated yield is 5.35 t/ha.
 E) To obtain an estimated yield of 5.2 t/ha, you need to apply 7.0 kg/ha of fertilizer.

**34)** Confidence Bands

The least squares regression line is the line:

 A) which is determined by use of a function of the distance between the observed Y ’s and the predicted Y’s.
 B) which has the smallest sum of the squared residuals of any line through the data values.
 C) for which the sum of the residuals about the line is zero.
 D) which has none of the above properties.
 E) which has all of the above properties.

**35)** Which diagnostic test allows the researcher to claim that her model explains x-percent of the variation in the dependent variable?

 A) Coefficient of Determination
 B) Durbin-Watson test
 C) t-test on slope coefficient
 D) Sum of squared residuals
 E) None of the options are correct.

**36)** The following regression represents seasonally adjusted shoe store sales in millions of dollars in the United States.

|  |
| --- |
| Audit Trail -- ANOVA Table (Multiple Regression Selected) |
| Source of variation | SS | df | MS | SEE |
| Regression |  | 2,099,394.07 |  |  | 1 |  |  | 2,099,394.07 |  |  |  |  |
| Error |  | 632,943.91 |  |  | 142 |  |  | 4,457.35 |  |  | 66.76 |  |
| Total |  | 2,732,337.98 |  |  | 143 |  |  |  |  |  |  |  |
|  |

|  |
| --- |
| Audit Trail -- Coefficient Table (Mulitple Regression Selected) |
| Series Description | Included in Model  | Coefficient  | Standard Error  | T-test  | P-value  | F-test  | Elasticity  | Overall F-test  |
| SASSS |  | Dependent |  |  | 1,523.74 |  |  | 11.19 |  |  | 136.23 |  |  | 0.00 |  |  | 18,557.38 |  |  |  |  |  | 471.00 |  |
| Time index |  | Yes |  |  | 2.90 |  |  | 0.13 |  |  | 21.70 |  |  | 0.00 |  |  | 471.00 |  |  | 0.12 |  |  |  |  |
|  |

|  |
| --- |
| Audit Trail - Statistics |
| Accuracy Measures | Value |  | Forecast Statistics | Value |
| AIC | 1,618.57 |  | Durbin Watson(12) | 1.15 |
| BIC | 1,621.54 |  | Mean | 1,734.33 |
| Mean Absolute Percentage Error (MAPE) | 3.06 | % | Standard Deviation | 138.23 |
| R-Square | 76.84 | % | Max | 1,988.06 |
| Mean Absolute Error | 51.97 |  | Min | 1,397.82 |
| Mean Square Error | 4,395.44 |  | Mean Absolute Deviation | 114.40 |
| Root Mean Square Error | 66.30 |  | Mean Square Deviation | 18,974.57 |
| Theil | 0.80 |  | Ljng-Box | 85.17 |
|  |

|  |  |
| --- | --- |
| Method Statistics | Value |
| Method Selected | Multiple Regression |
|  |

Consider the simple regression above. The dependent variable is seasonally adjusted shoe store sales in the United States. The independent variable is the index of time.In the shoe store sales regression above, the reported Durbin-Watson statistic

 A) indicates that the seasonal adjustment may not have eliminated all the seasonality.
 B) proves that trend does not exist in the shoe store sales data.
 C) indicates that the seasonal adjustment did eliminate any seasonal effect.
 D) indicates that there is no serial correlation.

**37)** Which of the following is a problem caused by heteroscedasticity when using ordinary least squares?

 A) The Y-intercept estimate is biased.
 B) The regression standard error estimate is biased.
 C) The slope estimate is biased.
 D) The slope estimate is unbiased.

**38)** Stock price data show periods of relatively calm interrupted by periods of enhanced price volatility. This suggests stock price data are

 A) autocorrelated.
 B) homoscedastic.
 C) linear.
 D) heteroscedastic.
 E) nonlinear.

**39)** If the plot of the residuals is fan shaped, which assumption of ordinary least squares regression is violated?

 A) No assumptions are violated, the graph should resemble a fan.
 B) Homoscedasticity
 C) Normality
 D) Independence of errors

**40)** Which of the following is not correct about causal regression analysis of the form Y = f(X)?

 A) Selection of the appropriate causal variable Y is important.
 B) Selection of the appropriate causal variable X is important.
 C) Use of economic theory to identify X is common.
 D) All of the options are correct.
 E) Use of past experience to identify X is common.

**41)** When we use the t distribution to evaluate the statistical significance of a regression coefficient,

 A) we are assuming that the distribution of the coefficient is binomial.
 B) we are calculating the R2.
 C) we want zero to be at the center of the distribution.
 D) we assume the distribution is centered on the calculated coefficient.

**42)** What is a common way forecasters attempt to eliminate heteroscedasticity?

 A) Transform the data using logarithms.
 B) Transform the data by first differencing.
 C) Apply Ordinary Least Squares.
 D) All of the options are correct.
 E) Deseasonalize the data.

**43)** How many critical regions does the Durbin-Watson test have?

 A) three
 B) four
 C) one
 D) two
 E) five

**44)** Consider the following model: Sales = α + β(TIME)2 + ε. The trend is modeled here as a

 A) exponential trend.
 B) quadratic trend.
 C) None of the options are correct.
 D) linear trend.
 E) logarithmic trend.

**45)** When severe autocorrelation is indicated after a regression model has been estimated, which underlying regression assumption is violated?

 A) The error terms are independent of each other.
 B) Heteroscedasticity
 C) The population of Y values is normally distributed about the population regression line.
 D) A linear relationship exists between X and Y in the population.
 E) The dispersion of population data points around the population regression line remains constant everywhere along the line.

**46)** Which of the following does not become unreliable when serial correlation is present?

 A) OLS slope estimates
 B) None of the options are correct.
 C) Error sum of squares
 D) R-squared
 E) t-tests

**47)** Table 4.2

|  |  |  |
| --- | --- | --- |
| Date | dpi | Time Index |
| Jan-93 | 4800.90 | 1 |
| Feb-93 | 4803.90 | 2 |
| Mar-93 | 4800.10 | 3 |
| Apr-93 | 4887.40 | 4 |
| May-93 | 4909.90 | 5 |
| Jun-93 | 4906.10 | 6 |
| Jul-93 | 4909.10 | 7 |
| Aug-93 | 4931.40 | 8 |
| Sep-93 | 4932.10 | 9 |
| Oct-93 | 4951.00 | 10 |
| Nov-93 | 4974.30 | 11 |
| Dec-93 | 5137.00 | 12 |
| Jan-94 | 4955.90 | 13 |
| Feb-94 | 5003.30 | 14 |
| Mar-94 | 5037.00 | 15 |
| Apr-94 | 5057.20 | 16 |
| May-94 | 5143.50 | 17 |
| Jun-94 | 5153.50 | 18 |
| Jul-94 | 5172.10 | 19 |
| Aug-94 | 5195.00 | 20 |
|  |

With the data in Table 4.2 from the text, a time trend regression was run.

 A) This regression was used as an example of "what not to do."
 B) This regression could be used to prove causality.
 C) This regression used the time index as an independent variable.
 D) This regression used the time index as a dependent variable.

**48)** Autocorrelation leads to or causes

 A) heteroscedasticity.
 B) nonlinear regression.
 C) serial correlation.
 D) spurious regression.
 E) All of the options are correct.

**49)** Fit and accuracy

 A) are the same things.
 B) do not depend on the estimated standard error.
 C) reflect in sample versus out-of-sample model forecast errors.
 D) All of the options are correct.
 E) do not depend on the fitted regression model.

**50)** X-Y data have been collected in which X ranges between 50 and 100 and Y ranges between 1200 and 2000. It is not wise to use the resulting regression line equation to predict Y when X is equal to −10 because

 A) a negative number cannot be used.
 B) All of the options are correct.
 C) the Y-intercept might be above zero.
 D) the proposed X value is well beyond the range of observed data.
 E) the predicted value for Y might turn out to be negative.

**51)** The expected trend value of September sales for a firm is $900. Assuming a September seasonal index of .91, what would be the seasonally adjusted forecast for September?

 A) $900
 B) None of the options are correct.
 C) $819
 D) $950
 E) $989

**52)** The sign on the slope estimate in a regression problem

 A) is the opposite of the sign of the correlation of Y and X.
 B) None of the options are correct.
 C) is the same as the sign of the Y-intercept.
 D) always has the same sign as the correlation of Y and X.
 E) has no relationship to the sign of the correlation of Y and X.

**53)** What possible decisions can be made using the Durbin-Watson test?

 A) Reject serial correlation.
 B) Conclude negative serial correlation.
 C) All of the options are correct.
 D) Conclude positive serial correlation.
 E) Conclude nothing.

**54)** Confidence BandsUse the confidence band diagram above.

 A) An approximate 95% confidence interval for a regression estimate would appear like lines EF and GH.
 B) An approximate 95% confidence interval for a regression estimate would appear like lines AB and CD.
 C) Neither set of lines (AB and CD nor EF and GH) represents a reasonable depiction of a confidence interval.
 D) Confidence intervals for least squares linear regression estimates may take on many shapes that appear quite unlike those depicted in the diagram.

**55)** The Y-intercept of a regression line is −14, and the slope is 3.5. Which of the following is not correct?

 A) The regression line crosses the Y-axis at −14.
 B) None of the options are correct.
 C) When X increases by one, Y increases by 3.5.
 D) X and Y are positively related.
 E) When Y increases by one, X increases by 3.5.

**56)** Income is used to predict savings. For the regression equation Y = 1,000 + .10X, which of the following is true?

 A) Y is income, X is savings, and savings is the independent variable.
 B) Y is income, X is savings, and income is the independent variable.
 C) Y is savings, X is income, and savings is the independent variable.
 D) Y is savings, X is income, and income is the independent variable.

**57)** The least squares procedure minimizes the sum of

 A) squared residuals.
 B) squared maximum error.
 C) None of the options are correct.
 D) the residuals.
 E) absolute errors.

**58)** Consider the following time trend regression model for explaining the behavior of disposable personal income (DPI): DPI = 17,000 + 41(TIME). If the regression standard error were 150, what is an approximate 95% prediction interval for quarter 3 DPI?

 A) None of the options are correct.
 B) 16,366 to 17,105
 C) 17,102 to 18,345
 D) 16,932 to 18,108
 E) 16,823 to 17,423

**59)** In a regression problem, if R-squared is equal to .90, it means that the ratio of the sum of

 A) squared Y values around the mean of Y to the sum of squared residuals is .90.
 B) squared residuals to the sum of squared Y values around the mean of Y is .90.
 C) squared Y values around the mean of Y to the sum of squared residuals is .10.
 D) squared residuals to the sum of squared Y values around the mean of Y is .10.
 E) None of the options are correct.

**60)** Which of the following is a tool used in model selection?

 A) Seasonality
 B) Cyclicity
 C) Plotting the data
 D) None of the options are correct.
 E) Growth

**61)** The regression slope term (β) in the simple bivariate regression model is

 A) undetermined using the OLS method.
 B) None of the options are correct.
 C) the change in the conditional mean of Y given a unit change in X.
 D) usually known to the investigator.
 E) correctly interpreted as dY/dX.

**62)** Confidence BandsConsider the diagram above representing the confidence bands around a linear least squares regression line.

 A) Neither set of lines nor the shaded area represent the true 95% confidence band.
 B) The lines AB and CD represent the true 95% confidence bands.
 C) The lines EF and GH represent the true 95% confidence bands.
 D) The shaded area represents the true 95% confidence band.

**63)** The strength of the linear relationship between two numerical variables may be measured by the

 A) scatter diagram.
 B) coefficient of determination.
 C) Y-intercept.
 D) slope.

**64)** Visual inspection of the data will help the forecaster identify

 A) seasonality.
 B) trend.
 C) linearity.
 D) nonlinearity.
 E) All of the options are correct.

**65)** The Standard Error of the Estimate (SEE) in an ordinary least squares regression

 A) is the standard error of the dependent variable.
 B) is the standard error of the independent variable.
 C) is used to test for autocorrelation.
 D) is used to test for serial correlation.

**66)** A residual is

 A) the difference between the mean of Y conditional on X and the unconditional mean.
 B) the difference between the mean of Y and its actual value.
 C) the difference between the sum of squared errors before and after X is used to predict Y.
 D) the difference between the regression prediction of Y and its actual value.
 E) None of the options are correct.

**67)** Cross-sectional regression models linking personal disposable income to consumption expenditure are likely to be hampered by

 A) homoscedasticity.
 B) None of the options are correct.
 C) nonlinearity.
 D) heteroscedasticity.
 E) autocorrelation.

**68)** In applying the Durbin-Watson test, the required inputs are

 A) All of the options are correct.
 B) number of independent variables.
 C) DW lower bound.
 D) sample size.
 E) estimated DW value.

**69)** In cross-sectional regression analysis,

 A) the time index is the only independent variable used.
 B) serial correlation is a serious problem.
 C) only a single independent variable may be used.
 D) the data all pertain to one time period.

**70)** In a linear least squares regression, what is one of the two parameters estimated?

 A) the coefficient of determination
 B) the dependent variable
 C) the independent variable
 D) the slope

**71)** Which of the following statements are true?

 A) None of the options are correct.
 B) Autocorrelation causes the sum of squares decomposition to become unreliable.
 C) Autocorrelation arises when there is a perfect linear association among the independent variables in the sample.
 D) Autocorrelation causes the ordinary least squares estimate of the error variance to become biased.
 E) Autocorrelation and its presence have no effect on the Gauss-Markov theorem.

**72)** When autocorrelation is present, which of the following is not a problem?

 A) The F-statistic may be unreliable.
 B) None of the options are correct.
 C) The t-statistics for each coefficient will be overstated.
 D) The R-squared statistic may be unreliable.
 E) The estimated standard errors will be larger than the true standard errors.

**73)** Which of the following is not true regarding testing for serial correlation?

 A) The DW statistic has a range from 0 to 4.
 B) The DW statistic has a range from 0 to 1.
 C) None of the options are correct.
 D) If negative serial correlation exists, the value of DW approaches 4.
 E) If positive serial correlation exists, the value of DW approaches zero.

**74)** The autocorrelation parameter is used to measure

 A) the slope of the regression line.
 B) disturbances or independent random variates.
 C) error or difference between a data point and the regression line.
 D) correlation between residuals.
 E) difference between present and past residuals.

**75)** Autocorrelation in a regression model occurs when there is some correlation

 A) between the error in one period and the error in the next period.
 B) None of the options are correct.
 C) among the explanatory variables.
 D) among the residuals and the values of the explanatory variables.
 E) between the slope and intercept estimates.

**76)**

|  |
| --- |
| Audit Trail -- Coefficient Table |
| Series Description | Included in Model | Coefficient | Standard Error | T-test | P-value |
| Jewelry Sales (Mil$) |  | Dependent |  | - | 71.571 |  |  | 555.841 |  | - | 0.13 |  |  | 0.90 |  |
| dpi |  | Yes |  |  | 0.29 |  |  | 0.09 |  |  | 3.46 |  |  | 0.00 |  |
|  |

|  |  |
| --- | --- |
| Audit Trail - Statistics |  |
| Accuracy Measures | Value |  | Forecast Statistics | Value |  |
| Mean Absolute Percentage Error (MAPS) | 27.07 | % | Durbin Watson(1) | 1.96 |  |
| R-Square | 8.45 | % | Mean | 1,828.48 |  |
| Root Mean Square Error | 1,035.87 |  |  |  |  |
|  |

The jewelry regression shown

 A) is a seasonal regression.
 B) is a causal regression.
 C) is a linear time trend regression.
 D) is a nonlinear regression.

**77)** If the correlation coefficient in an ordinary least squares regression = 1.00, then

 A) all the data points must fall exactly on a horizontal straight line with a zero slope.
 B) all the data points must fall exactly on a straight line with a positive slope.
 C) all the data points must fall exactly on a straight line with a negative slope.
 D) all the data points must fall exactly on a straight line with a slope that equals 1.00.

**78)** Regression model disturbances (forecast errors)

 A) All of the options are correct.
 B) are assumed to average to zero.
 C) are assumed to follow a normal probability distribution.
 D) are assumed to be independent over time.
 E) can be estimated by OLS residuals.

**79)** The following regression represents seasonally adjusted shoe store sales in millions of dollars in the United States.

|  |
| --- |
| Audit Trail -- ANOVA Table (Multiple Regression Selected) |
| Source of variation | SS | df | MS | SEE |
| Regression |  | 2,099,394.07 |  |  | 1 |  |  | 2,099,394.07 |  |  |  |  |
| Error |  | 632,943.91 |  |  | 142 |  |  | 4,457.35 |  |  | 66.76 |  |
| Total |  | 2,732,337.98 |  |  | 143 |  |  |  |  |  |  |  |
|  |

|  |
| --- |
| Audit Trail -- Coefficient Table (Mulitple Regression Selected) |
| Series Description | Included in Model  | Coefficient  | Standard Error  | T-test  | P-value  | F-test  | Elasticity  | Overall F-test  |
| SASSS |  | Dependent |  |  | 1,523.74 |  |  | 11.19 |  |  | 136.23 |  |  | 0.00 |  |  | 18,557.38 |  |  |  |  |  | 471.00 |  |
| Time index |  | Yes |  |  | 2.90 |  |  | 0.13 |  |  | 21.70 |  |  | 0.00 |  |  | 471.00 |  |  | 0.12 |  |  |  |  |
|  |

|  |
| --- |
| Audit Trail - Statistics |
| Accuracy Measures | Value |  | Forecast Statistics | Value |
| AIC | 1,618.57 |  | Durbin Watson(12) | 1.15 |
| BIC | 1,621.54 |  | Mean | 1,734.33 |
| Mean Absolute Percentage Error (MAPE) | 3.06 | % | Standard Deviation | 138.23 |
| R-Square | 76.84 | % | Max | 1,988.06 |
| Mean Absolute Error | 51.97 |  | Min | 1,397.82 |
| Mean Square Error | 4,395.44 |  | Mean Absolute Deviation | 114.40 |
| Root Mean Square Error | 66.30 |  | Mean Square Deviation | 18,974.57 |
| Theil | 0.80 |  | Ljng-Box | 85.17 |
|  |

|  |  |
| --- | --- |
| Method Statistics | Value |
| Method Selected | Multiple Regression |
|  |

Consider the simple regression above. The dependent variable is seasonally adjusted shoe store sales in the United States. The independent variable is the index of time.The Durbin Watson (12) statistic is reported because

 A) the researcher is examining multicollinearity.
 B) the time series was probably monthly data.
 C) the data has 12 observations available for analysis.
 D) there are 12 degrees of freedom available for calculating the significance of the t-statistic.

**80)** The Y-intercept (b0) in an ordinary least squares regression represents the

 A) variation around the sample regression line.
 B) change in the estimated average Y per unit change in X.
 C) predicted value of Y when X=0.
 D) predicted value of Y.

**81)** The least squares procedure minimizes the sum of

 A) None of the options are correct.
 B) absolute errors.
 C) squared residuals.
 D) squared maximum error.
 E) the residuals.

**82)** Seasonal indices of sales for the Black Lab Ski Resort are for 1.20 for January and .80 for December. If December sales for 1998 were $5,000, a reasonable estimate of sales for January 1999 is

 A) $6,000.
 B) $10,000.
 C) $7,500.
 D) None of the options are correct.
 E) $4,800.

**83)** For children, there is approximately a linear relationship between "height" and "age". One child was measured monthly. Her height was 75 cm at 3 years of age and 85 cm when she was measured 18 months later. A least squares line was fit to her data. The slope of this line is approximately:

 A) 10 cm/m.
 B) 1.57 cm/m.
 C) 25 cm/m.
 D) 2.1 cm/m.
 E) 0.55 cm/m.

**84)** The following output resulted from a regression model where SAGap is seasonally adjusted Gap sales and dpi is disposable income per capita.

|  |
| --- |
| Audit Trail -- Coefficient Table (Mulitple Regression Selected) |
| Series Description | Included in Model  | Coefficient  | Standard Error  | T-test  | P-value  | F-test  | Elasticity  |
| SAGAP |  | Dependent |  | - | 2,867,564.78 |  |  | 140,536.33 |  | - | 20.40 |  |  | 0.00 |  |  | 416.34 |  |  |  |  |
| dpi |  | Yes |  |  | 809.79 |  |  | 25.04 |  |  | 32.33 |  |  | 0.00 |  |  | 1,045.55 |  |  | 2.91 |  |
|  |

|  |
| --- |
| Audit Trail -- Correlation Coefficient Table |
| Series Description | Included in Model | SAGap | dpi |
| SAGap |  | Dependent |  |  | 1.00 |  |  | 0.97 |  |
| dpi |  | Yes |  |  | 0.97 |  |  | 1.00 |  |
|  |

|  |  |
| --- | --- |
| Audit Trail - Statistics |  |
| Accuracy Measures | Value |  | Forecast Statistics | Value |  |
| AIC | 2,135.23 |  | Durbin Watson(4) | 0.43 |  |
| BIC | 2,137.56 |  | Mean | 1,501,041.15 |  |
| Mean Absolute Percentage Error (MAPE) | 34.60 | % | Standard Deviation | 1,303,264.45 |  |
| R-Square | 94.59 | % | Max | 4,253,174.95 |  |
| Adjusted R-Square | 94.51 | % | Min | 123,121.77 |  |
| Mean Square Error | 90,711,613,878.48 |  | Range | 4,130,053.18 |  |
| Root Mean Square Error | 301,183.69 |  | Root Mean Square | 1,294,661.95 |  |
| Theil | 6.23 |  | Ljng-Box | 347.40 |  |
|  |  |

Linear least squares regression chooses values for the intercept and slope to minimize

 A) None of the options are correct.
 B) the sum of the squared errors (distances) from the average value of the dependent variable.
 C) the sum of the squared errors (distances) from the regression line.
 D) the sum of the squared errors (distances) from the reference line.
 E) the R2 (coefficient of determination) of the regression.

**85)** The most common mathematical trend equation for a time series is called the least squares trend because it is the line which minimizes the sum of the

 A) squared vertical deviations from the trend line.
 B) deviations from the mean of the X variable.
 C) deviations from the mean.
 D) None of the options are correct.
 E) squares of deviations from the sample mean.