Forecasting and Predictive Analytics with Forecast X, 7e **(Keating)**
Chapter 2 The Forecast Process, Data Considerations, and Model Selection

1) Why are forecasting textbooks full of applied statistics?
A) Statistics is the study of uncertainty.
B) Real-world business decisions involve risk and uncertainty.
C) Forecasting attempts to generate certainty out of uncertain events.
D) Forecasting ultimately deals with probability.
E) All of the options are correct.

Answer: E
Difficulty: 1 Easy
Topic: The Forecast Process
Learning Objective: 2-01 Explain a process for developing forecasts.
Accessibility: Keyboard Navigation
Gradable: automatic

2) Which of the following is not part of the recommended nine-step forecast process?
A) What role do forecasts play in the business decision process?
B) What exactly is to be forecast?
C) How urgent is the forecast?
D) Is there enough data?
E) All of the options are correct.

Answer: E
Difficulty: 1 Easy
Topic: The Forecast Process
Learning Objective: 2-01 Explain a process for developing forecasts.
Accessibility: Keyboard Navigation
Gradable: automatic

3) Of the following model selection criteria, which is often the most important in determining the appropriate forecast method?
A) Technical background of the forecast user
B) Patterns the data have exhibited in the past
C) How much money is in the forecast budget?
D) What is the forecast horizon?
E) When is the forecast needed?

Answer: B
Difficulty: 1 Easy
Topic: Data Patterns and Model Selection
Learning Objective: 2-03 Identify forecasting methods that would be good candidates for a given series to be forecast.
Accessibility: Keyboard Navigation
Gradable: automatic

4) Time series data of a typical The GAP store should show which of the following data patterns?
A) Trend
B) Seasonal
C) Cyclical
D) Random
E) All of the options are correct.

Answer: E
Difficulty: 1 Easy
Topic: Trend, Seasonal and Cyclical Data Patterns
Learning Objective: 2-02 Distinguish between trend, seasonal, and cyclical data patterns.
Accessibility: Keyboard Navigation
Gradable: automatic

5) Which of the following is incorrect?
A) The forecaster should be able to defend why a particular model or procedure has been chosen.
B) Forecast errors should be discussed in an objective manner to maximize management's confidence in the forecast process.
C) Forecast errors should not be discussed since most people know that forecasting is an inexact science.
D) You should tailor your presentation to the sophistication of the audience to maximize credibility in the forecast process.
E) None of the options are correct.

Answer: C
Difficulty: 1 Easy
Topic: The Forecast Process
Learning Objective: 2-03 Identify forecasting methods that would be good candidates for a given series to be forecast.
Accessibility: Keyboard Navigation
Gradable: automatic

6) In the model-testing phase of the nine-step process, which of the following refers to that portion of a sample used to evaluate model-forecast accuracy?
A) Fit
B) Forecast horizon
C) Holdout period
D) Accuracy
E) None of the options are correct.

Answer: C
Difficulty: 1 Easy
Topic: The Forecast Process
Learning Objective: 2-01 Explain a process for developing forecasts.
Accessibility: Keyboard Navigation
Gradable: automatic

7) Your authors present a guide to selecting an appropriate forecasting method based on
A) data patterns.
B) quantity of historical data available.
C) forecast horizon.
D) quantitative background of the forecast user.
E) All of the options are correct.

Answer: E
Difficulty: 1 Easy
Topic: The Forecast Process
Learning Objective: 2-01 Explain a process for developing forecasts.
Accessibility: Keyboard Navigation
Gradable: automatic

8) Which time-series component is said to fluctuate around the long-term trend and is fairly irregular in appearance?
A) Trend.
B) Cyclical.
C) Seasonal.
D) Irregular.
E) None of the options are correct.

Answer: B
Difficulty: 1 Easy
Topic: Trend, Seasonal and Cyclical Data Patterns
Learning Objective: 2-02 Distinguish between trend, seasonal, and cyclical data patterns.
Accessibility: Keyboard Navigation
Gradable: automatic

9) Forecasting January sales based on the previous month's level of sales is likely to lead to error if the data are _____.
A) stationary
B) non-cyclical
C) seasonal
D) irregular
E) None of the options are correct.

Answer: C
Difficulty: 1 Easy
Topic: Trend, Seasonal and Cyclical Data Patterns
Learning Objective: 2-02 Distinguish between trend, seasonal, and cyclical data patterns.
Accessibility: Keyboard Navigation
Gradable: automatic

10) The difference between seasonal and cyclical components is
A) duration
B) source
C) predictability
D) frequency
E) All of the options are correct.

Answer: E
Difficulty: 1 Easy
Topic: Trend, Seasonal and Cyclical Data Patterns
Learning Objective: 2-02 Distinguish between trend, seasonal, and cyclical data patterns.
Accessibility: Keyboard Navigation
Gradable: automatic

11) For which data frequency is seasonality not a problem?
A) Daily.
B) Weekly.
C) Monthly.
D) Quarterly.
E) Annual.

Answer: E
Difficulty: 1 Easy
Topic: Trend, Seasonal and Cyclical Data Patterns
Learning Objective: 2-02 Distinguish between trend, seasonal, and cyclical data patterns.
Accessibility: Keyboard Navigation
Gradable: automatic

12) One can realistically not expect to find a model that fits any data set perfectly due to the _____ component of a time series.
A) Trend
B) Seasonal
C) Cyclical
D) Irregular
E) None of the options are correct.

Answer: D
Difficulty: 1 Easy
Topic: Trend, Seasonal and Cyclical Data Patterns
Learning Objective: 2-02 Distinguish between trend, seasonal, and cyclical data patterns.
Accessibility: Keyboard Navigation
Gradable: automatic

13) When a time series contains no trend, it is said to be
A) nonstationary.
B) seasonal.
C) nonseasonal.
D) stationary.
E) filtered.

Answer: D
Difficulty: 1 Easy
Topic: Trend, Seasonal and Cyclical Data Patterns
Learning Objective: 2-02 Distinguish between trend, seasonal, and cyclical data patterns.
Accessibility: Keyboard Navigation
Gradable: automatic

14) Stationarity refers to
A) the size of the RMSE of a forecasting model.
B) the size of variances of the model's estimates.
C) a method of forecast optimization.
D) lack of trend in a given time series.
E) None of the options are correct.

Answer: D
Difficulty: 1 Easy
Topic: Trend, Seasonal and Cyclical Data Patterns
Learning Objective: 2-02 Distinguish between trend, seasonal, and cyclical data patterns.
Accessibility: Keyboard Navigation
Gradable: automatic

15) Which of the following is not a measure of central tendency in a population?
A) Mean.
B) Mode.
C) Median.
D) Range.

Answer: D
Difficulty: 1 Easy
Topic: A Statistical Review
Learning Objective: 2-04 Explain the differences between the mean, median, and mode for a set of data.
Accessibility: Keyboard Navigation
Gradable: automatic

16) Which of the following is not a descriptive statistic?
A) Expected value.
B) Mean.
C) Range.
D) Variance.
E) None of the options are correct.

Answer: E
Difficulty: 1 Easy
Topic: Descriptive Statistics
Learning Objective: 2-04 Explain the differences between the mean, median, and mode for a set of data.
Accessibility: Keyboard Navigation
Gradable: automatic

17) Which of the following is not a foundation of classical statistics?
A) Summary measures of probability distributions called descriptive statistics
B) Probability distribution functions, which characterize all outcomes of a variable
C) The use of sampling distributions, which describe the uncertainty in making inference about the population on the basis of a sample
D) The concept of expected value
E) None of the options are correct.

Answer: E
Difficulty: 2 Medium
Topic: Descriptive Statistics
Learning Objective: 2-04 Explain the differences between the mean, median, and mode for a set of data.
Accessibility: Keyboard Navigation
Gradable: automatic

18) The standard normal probability table
A) is equivalent to a t distribution if the sample size is less than 30.
B) shows a normal distribution with standard deviation equal to zero.
C) is used to make inference for all normally distributed random variables.
D) All of the options are correct.
E) None of the options are correct.

Answer: C
Difficulty: 1 Easy
Topic: The Normal Distribution
Learning Objective: 2-05 Explain the most common measures of dispersion in data.
Accessibility: Keyboard Navigation
Gradable: automatic

19) The median and mode may be more accurate than the sample mean in forecasting the populations mean when
A) the sample size is small.
B) the sample size is large.
C) the sample has one large outlier.
D) the population is assumed to be normally distributed.
E) All of the options are correct.

Answer: C
Difficulty: 1 Easy
Topic: Descriptive Statistics
Learning Objective: 2-04 Explain the differences between the mean, median, and mode for a set of data.
Accessibility: Keyboard Navigation
Gradable: automatic

20) The arithmetic average of the relative frequency of the occurrence of some random variable is also called the _____.
A) range
B) mean
C) variance
D) standard deviation
E) None of the options are correct.

Answer: B
Difficulty: 1 Easy
Topic: Descriptive Statistics
Learning Objective: 2-04 Explain the differences between the mean, median, and mode for a set of data.
Accessibility: Keyboard Navigation
Gradable: automatic

21) In finance, an investor who ignores risk is termed "risk neutral." What descriptive statistic is our risk neutral investor ignoring when she generates stock portfolios?
A) Median.
B) Mean.
C) Mode.
D) Standard deviation.
E) None of the options are correct.

Answer: D
Difficulty: 3 Hard
Topic: Descriptive Statistics
Learning Objective: 2-04 Explain the differences between the mean, median, and mode for a set of data.
Accessibility: Keyboard Navigation
Gradable: automatic

22) In calculating the sample variance, we subtract one from the sample size. This is because
A) the population mean is unknown.
B) of using the sample mean to estimate the population mean.
C) the sum of deviations about the sample mean is zero.
D) the sample mean is employed.
E) All of the options are correct.

Answer: E
Difficulty: 3 Hard
Topic: Descriptive Statistics
Learning Objective: 2-04 Explain the differences between the mean, median, and mode for a set of data.
Accessibility: Keyboard Navigation
Gradable: automatic

23) Which statistic is correctly interpreted as the "average" spread of data about the mean?
A) Mode.
B) Range.
C) Variance.
D) Standard deviation.
E) Mean.

Answer: D
Difficulty: 1 Easy
Topic: Descriptive Statistics
Learning Objective: 2-04 Explain the differences between the mean, median, and mode for a set of data.
Accessibility: Keyboard Navigation
Gradable: automatic

24) Which measure of dispersion in a data set is the most intuitive and represents an average?
A) Range.
B) Mode.
C) Standard deviation.
D) Variance.
E) Mean.

Answer: C
Difficulty: 1 Easy
Topic: Descriptive Statistics
Learning Objective: 2-04 Explain the differences between the mean, median, and mode for a set of data.
Accessibility: Keyboard Navigation
Gradable: automatic

25) Which of the following is not an attribute of a normal probability distribution?
A) It is symmetrical about the mean.
B) Most observations cluster around the mean.
C) Most observations cluster around zero.
D) The distribution is completely determined by the mean and variance.
E) All of the options are correct.

Answer: C
Difficulty: 1 Easy
Topic: The Normal Distribution
Learning Objective: 2-06 Discuss the normal and Student's t distributions.
Accessibility: Keyboard Navigation
Gradable: automatic

26) Which of the following is not a foundation of classical statistics?
A) Summary measures of probability distribution called descriptive statistics.
B) Probability distribution function which characterizes all possible outcomes of a random variable.
C) The knowledge of thousands and thousands of normal probability tables required for statistical inference of normally distributed random variables.
D) The concept of expected value, which is the average value of a random variable taken over a large number of samples.

Answer: C
Difficulty: 3 Hard
Topic: A Statistical Review
Learning Objective: 2-06 Discuss the normal and Student's t distributions.
Accessibility: Keyboard Navigation
Gradable: automatic

27) A company claims that the rubber belts, which it manufactures, have a mean service life of at least 800 hours. A random sample of 36 belts from a very large shipment of the company's belts shows a mean life of 760 hours and a standard deviation of 90 hours. Which of the following is the most appropriate on the basis of the sample results?
A) The sample results do not warrant rejection of the company's claim if the risk of a Type I error is specified at .05.
B) The sample results do warrant rejection of the company's claim if the risk of Type I error is specified at .05.
C) Since the sample mean falls below the company's claim, the sample results indicate that the company claim is incorrect.
D) The sample results are indeterminate since the magnitude of the sample standard deviation is greater than the difference between the company's claimed figure and the sample mean.

Answer: B
Difficulty: 1 Easy
Topic: Hypothesis Testing
Learning Objective: 2-07 Describe three common forms of statistical hypotheses.
Accessibility: Keyboard Navigation
Gradable: automatic

28) Based upon ten years of monthly data, the monthly rate of return of the DOW Jones 30 composite stock portfolio was normally distributed with mean .0084 and variance .0014. What is the probability, that in any given month, we observe a rate of return on the DOW above 10 percent?
A) Less than one percent.
B) Two percent.
C) Three percent.
D) Not enough information is provided to answer the question.

Answer: A
Difficulty: 2 Medium
Topic: A Statistical Review
Learning Objective: 2-06 Discuss the normal and Student's t distributions.
Accessibility: Keyboard Navigation
Gradable: automatic

29) Suppose you observe the entire population of a random variable and you wish to test some hypothesis about the mean. To perform your hypothesis test, you
A) apply a sampling distribution to the problem.
B) obtain sample estimates of population parameters.
C) simply find the population mean and compare it to the hypothesized value.
D) apply the t distribution.
E) There is no answer to this question.

Answer: C
Difficulty: 1 Easy
Topic: From Sample to Population: Statistical Inference
Learning Objective: 2-07 Describe three common forms of statistical hypotheses.
Accessibility: Keyboard Navigation
Gradable: automatic

30) If two large random samples are drawn from two populations, each having a mean of $100, the relevant sampling distribution of their difference has a mean of
A) $200.
B) the sum of the two sample means.
C) 0.
D) the difference between the two sample means.

Answer: C
Difficulty: 1 Easy
Topic: Descriptive Statistics
Learning Objective: 2-04 Explain the differences between the mean, median, and mode for a set of data.
Accessibility: Keyboard Navigation
Gradable: automatic

31) Type I error
A) is said to arise when we reject a true null hypothesis.
B) has a probability value equal to the significance level of any statistical test.
C) is a measure of the uncertainty associated with rejecting any null hypothesis on the basis of sample data.
D) Both A and B are correct.
E) All of the options are correct.

Answer: E
Difficulty: 1 Easy
Topic: Hypothesis Testing
Learning Objective: 2-07 Describe three common forms of statistical hypotheses.
Accessibility: Keyboard Navigation
Gradable: automatic

32) Sampling distributions
A) are the distributions of all possible values of a sample statistic based upon repeated sampling.
B) are used to make inference when the population of a variable is unobservable.
C) exhibit important properties for the ranking of alternative estimators such as unbiasedness and efficiency.
D) All of the options are correct.

Answer: D
Difficulty: 1 Easy
Topic: From Sample to Population: Statistical Inference
Learning Objective: 2-06 Discuss the normal and Student's t distributions.
Accessibility: Keyboard Navigation
Gradable: automatic

33) An unbiased model
A) is one that does not consistently over-estimate or under-estimate the true value of a parameter.
B) is one that consistently produces estimates with the smallest RMSE.
C) is one which contains no independent variable; it depends solely on time-series pattern recognition.
D) is one made up by a team of forecasters.

Answer: A
Difficulty: 2 Medium
Topic: Data Patterns and Model Selection
Learning Objective: 2-03 Identify forecasting methods that would be good candidates for a given series to be forecast.
Accessibility: Keyboard Navigation
Gradable: automatic

34) Suppose that you mistakenly move the decimal point to the right one digit in data from a normal population with a mean of zero. What happens to the standard deviation?
A) Data with mistake has standard error ten times the original.
B) Data with mistake has same standard error as the original.
C) Data with the mistake has twice the standard error of the original.
D) Data with the mistake has one hundred times the standard error of the original.
E) None of the options are correct.

Answer: A
Difficulty: 3 Hard
Topic: A Statistical Review
Learning Objective: 2-05 Explain the most common measures of dispersion in data.
Accessibility: Keyboard Navigation
Gradable: automatic

35) Which statement is incorrect?
A) Confidence intervals depend on sample size.
B) The sample mean is the best estimator if sampling from a normal population.
C) The sample mean is an unbiased estimator.
D) Confidence intervals provide no more information than point estimates.
E) The sample variance is an unbiased estimator.

Answer: D
Difficulty: 2 Medium
Topic: A Statistical Review
Learning Objective: 2-05 Explain the most common measures of dispersion in data.
Accessibility: Keyboard Navigation
Gradable: automatic

36) A machine fills ketchup bottles. One of the requirements is that the mean content of the bottles should be 10 ounces. Management wishes to set up a decision rule to decide whether or not this is true based on a random sample of bottles. The risk of type I error is specified at .05. A sample of 100 bottles will be taken; it is believed that the standard deviation of fills is .3 ounces. If it is decided that $Z = 2$, the decision rule boundary values are
A) 9.60 and 10.40.
B) 9.10 and 9.90.
C) 9.94 and 10.05.
D) 9.40 and 10.60.
E) None of the options are correct.

Answer: C
Difficulty: 3 Hard
Topic: A Statistical Review
Learning Objective: 2-05 Explain the most common measures of dispersion in data.
Accessibility: Keyboard Navigation
Gradable: automatic

37) Last year's midterm results showed a mean of 51 points and a variance of 46. An approximate confidence interval is
A) 44.2 to 57.8.
B) 37.4 to 64.6.
C) 5 to 97.
D) None of the options are correct.

Answer: B
Difficulty: 2 Medium
Topic: From Sample to Population: Statistical Inference
Learning Objective: 2-07 Describe three common forms of statistical hypotheses.
Accessibility: Keyboard Navigation
Gradable: automatic

38) A difference between the population standard deviation of the random variable X and the standard deviation of the sampling distribution of the sample mean is
A) one is based upon the other.
B) dependence on sample size.
C) the possibility of sampling error.
D) application to the t distribution.
E) All of the options are correct.

Answer: E
Difficulty: 3 Hard
Topic: From Sample to Population: Statistical Inference
Learning Objective: 2-07 Describe three common forms of statistical hypotheses.
Accessibility: Keyboard Navigation
Gradable: automatic

39) Which probability distribution is appropriate for testing hypotheses concerning an unknown population mean when the sample variance is used to estimate the population variance?
A) The normal distribution with mean μ and variance σ^2.
B) The normal distribution with mean 0 and variance 1.
C) The standard normal distribution.
D) The t distribution with n-1 degrees of freedom.
E) None of the options are correct.

Answer: D
Difficulty: 1 Easy
Topic: The Student's t-Distribution
Learning Objective: 2-06 Discuss the normal and Student's t distributions.
Accessibility: Keyboard Navigation
Gradable: automatic

40) A random sample of bolts is taken from inventory, and their lengths are measured. The average length in the sample is 5.3 inches with a standard deviation of .2 inches. The sample size was 50. The point estimate for the mean length of all bolts in inventory is
A) 5.3 inches.
B) .2 inches.
C) 4.908 to 5.692 inches.
D) 5.3 inches plus or minus .2

Answer: A
Difficulty: 2 Medium
Topic: From Sample to Population: Statistical Inference
Learning Objective: 2-07 Describe three common forms of statistical hypotheses.
Accessibility: Keyboard Navigation
Gradable: automatic

41) A random sample of bolts is taken from inventory, and their lengths are measured. The average length in the sample is 5.3 inches with a standard deviation of .2 inches. The sample size was 50.

A 95% confidence interval for the unknown population mean is
A) 5.3 inches.
B) 4.9 to 5.7 inches.
C) 5.3 inches plus or minus .056.
D) 4.784 to 5.816 inches.
E) None of the options are correct.

Answer: C
Difficulty: 2 Medium
Topic: From Sample to Population: Statistical Inference
Learning Objective: 2-07 Describe three common forms of statistical hypotheses.
Accessibility: Keyboard Navigation
Gradable: automatic

42) Which of the following statements about the probability of Type I and Type II error is not correct?
A) Type I error cannot occur if the null hypothesis is false.
B) Type II error cannot occur if the null hypothesis is true.
C) If the null hypothesis is true, the results of the test will either be a correct conclusion or a Type I error.
D) It is not possible to specify both the probabilities of Type I and II errors since only one of them can occur.

Answer: D
Difficulty: 2 Medium
Topic: Hypothesis Testing
Learning Objective: 2-07 Describe three common forms of statistical hypotheses.
Accessibility: Keyboard Navigation
Gradable: automatic

43) A sample of 100 are selected at random from a process with a mean of 500.52 and a standard deviation of 4.0. Estimate the probability that a sample of 100 would have a mean equal to or greater than 500.52 if the true population mean is really 500.0.
A) about .4938.
B) about .9032.
C) about .0968.
D) about .4032.

Answer: C
Difficulty: 2 Medium
Topic: From Sample to Population: Statistical Inference
Learning Objective: 2-07 Describe three common forms of statistical hypotheses.
Accessibility: Keyboard Navigation
Gradable: automatic

44) A random sample of employee files is drawn revealing an average of 2.8 overtime hours worked per week with a standard deviation of .7; the sample size is 500. The resulting 90% confidence interval is
A) 2.1 to 3.5.
B) 2.6 to 3.5.
C) 2.75 o 2.85.
D) 2.6 to 3.0.
E) None of the options are correct.

Answer: C
Difficulty: 2 Medium
Topic: From Sample to Population: Statistical Inference
Learning Objective: 2-07 Describe three common forms of statistical hypotheses.
Accessibility: Keyboard Navigation
Gradable: automatic

45) A medical researcher has just calculated a correlation coefficient of zero for two particular random variables. Which of the following statements is most accurate?
A) There is no significant linear difference between the two variables.
B) There is no significant relationship between the two variables.
C) There is no significant linear relationship between the two variables.
D) There is a significant linear relationship between the two variables.

Answer: C
Difficulty: 1 Easy
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

46) The correlation coefficient (ρ) is an extremely important descriptive statistic because
A) It provides a unit-free measure of how two random variables move together.
B) It provides a measure of the linear association between a pair of random variables.
C) It provides the forecaster with a diagnostic tool of when regression analysis is appropriate for the business-forecasting problem.
D) All of the options are correct.

Answer: D
Difficulty: 1 Easy
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

47) A large sample of X-Y data values are analyzed and reveal a correlation coefficient of −.88. Which statement is correct?
A) If r had been +.88, the correlation would have been much stronger.
B) The correlation is weak because r is less than -1.
C) A fairly strong negative linear relationship exists.
D) A weak negative relationship exists.

Answer: C
Difficulty: 1 Easy
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

48) Suppose two random variables X and Y are related as follows: $Y = 1/X^2$. The population Pearson correlation coefficient should be
A) +1.
B) 0.
C) −1.
D) .5.
E) None of the options are correct.

Answer: B
Difficulty: 2 Medium
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

49) Which functions are not appropriate for use of the Pearson correlation coefficient to estimate the correlation between a pair of random variables?
A) Cubic polynomials.
B) Quadratic polynomials.
C) Higher-order polynomials.
D) Functions involving a variable raised to the one-half power.
E) Reciprocal functions.

Answer: A
Difficulty: 1 Easy
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

50) If we were to know the true population correlation, confidence intervals for the population correlation can be constructed using the _____ distribution.
A) t distribution
B) standard normal distribution
C) chi-square distribution
D) F distribution
E) All of the options are correct.

Answer: B
Difficulty: 2 Medium
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

51) If the scatterplot of two variables has a circular pattern, this suggests the two variables have a population correlation coefficient of
A) −1.
B) −.5.
C) 0.
D) +.5.
E) +1.

Answer: C
Difficulty: 1 Easy
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

52) Which of the following is not used to calculate the sample Pearson correlation coefficient for the variables X and Y?
A) Sample mean of X.
B) Sample mean of Y.
C) Sample covariance of X and Y.
D) Sample standard deviation of X.
E) All of the options are used to calculate correlation coefficients.

Answer: E
Difficulty: 1 Easy
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

53) Which of the following is not a benefit of a scatter diagram?
A) The nature of the X-Y relationship (linear of nonlinear) may be revealed.
B) The strength of the relationship may be revealed.
C) The sign of the correlation coefficient will be revealed.
D) Displaying the population size.

Answer: D
Difficulty: 1 Easy
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

54) In order to conduct a correlation analysis, the collected data must be
A) related to the real world.
B) numerical.
C) constructed of categories.
D) highly correlated.
E) All of the options are correct.

Answer: B
Difficulty: 1 Easy
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

55) Which of the following is not a reason for testing if the population correlation coefficient is zero?
A) To see if r and rho (ρ. are equal)
B) To make inference from sample to population.
C) To bring sample size into the analysis.
D) To determine if a significant X-Y relationship exists.
E) All of the options are correct.

Answer: A
Difficulty: 1 Easy
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

56) Suppose the sample Pearson correlation coefficient (r) is estimated to be .75 with a sample size of 35. The correct calculated value of the test statistic for a null of zero correlation is
A) 6.5.
B) 8.6.
C) 44.1.
D) 2.5
E) None of the options are correct.

Answer: A
Difficulty: 1 Easy
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

57) For a collection of 15 X-Y data values, the sample correlation coefficient was estimated at -.63 from a sample of size 15. The calculated t value for a null of zero correlation is
A) 2.92.
B) 1.92.
C) −2.92.
D) −1.92.
E) None of the options are correct.

Answer: C
Difficulty: 1 Easy
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

58) When the correlation coefficient is negative, it means
A) there is a weak relationship.
B) when X goes down, Y does too.
C) X will not be a good predictor of Y.
D) when X goes down, Y tends to go up.
E) None of the above.

Answer: D
Difficulty: 1 Easy
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

59) When forecasting with time-series data, it is highly recommended to test for the presence of a trend in the data. Testing for trend at the 10% level of significance
A) can be accomplished by use of a standard 95% correlogram.
B) requires use of the standard normal probability distribution.
C) can be accomplished by comparing the estimated autocorrelation coefficient with the number 2 divided by the square root of sample size.
D) requires use of the t distribution.

Answer: D
Difficulty: 2 Medium
Topic: Correlograms
Learning Objective: 2-09 Explain how an autocorrelation function (ACF) can be useful in forecasting.
Accessibility: Keyboard Navigation
Gradable: automatic

60) Quarterly time-series data with a trend can be applied to models that assume stationary data by
A) averaging the data over time.
B) taking the first difference of the original series.
C) taking the fourth difference of the original series.
D) using a moving average.

Answer: B
Difficulty: 1 Easy
Topic: Correlograms
Learning Objective: 2-09 Explain how an autocorrelation function (ACF) can be useful in forecasting.
Accessibility: Keyboard Navigation
Gradable: automatic

61) Which of the following is not consistent with the presence of a trend in a time series?
A) The autocorrelation function declines quickly to zero as the lag increases.
B) The autocorrelation function of the first-differences declines quickly to zero as the lag increases.
C) The autocorrelation function declines slowly towards zero as the lag increases.
D) The autocorrelation function of the first-differences quickly declines to zero.

Answer: A
Difficulty: 2 Medium
Topic: Correlograms
Learning Objective: 2-09 Explain how an autocorrelation function (ACF) can be useful in forecasting.
Accessibility: Keyboard Navigation
Gradable: automatic

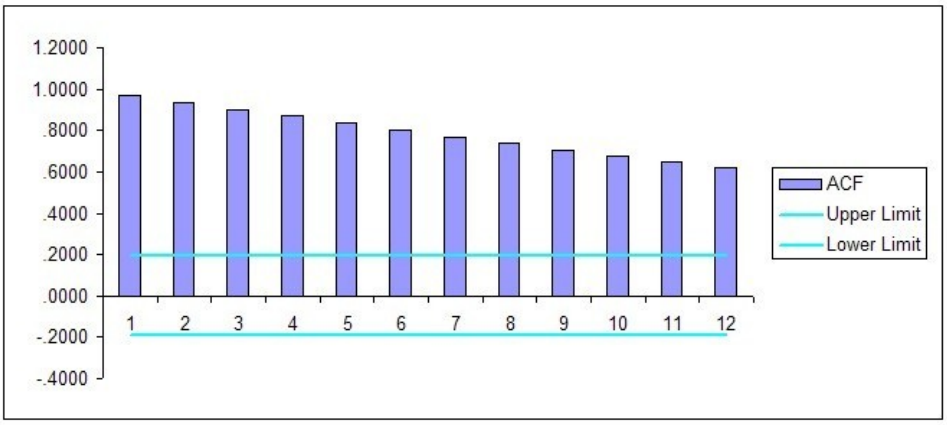
62) Autocorrelation refers to the correlation between a variable and
A) itself.
B) another very similar variable.
C) itself when lagged one or more periods.
D) another variable when the analysis is done on a computer.
E) None of the options are correct.

Answer: C
Difficulty: 1 Easy
Topic: Correlograms
Learning Objective: 2-09 Explain how an autocorrelation function (ACF) can be useful in forecasting.
Accessibility: Keyboard Navigation
Gradable: automatic

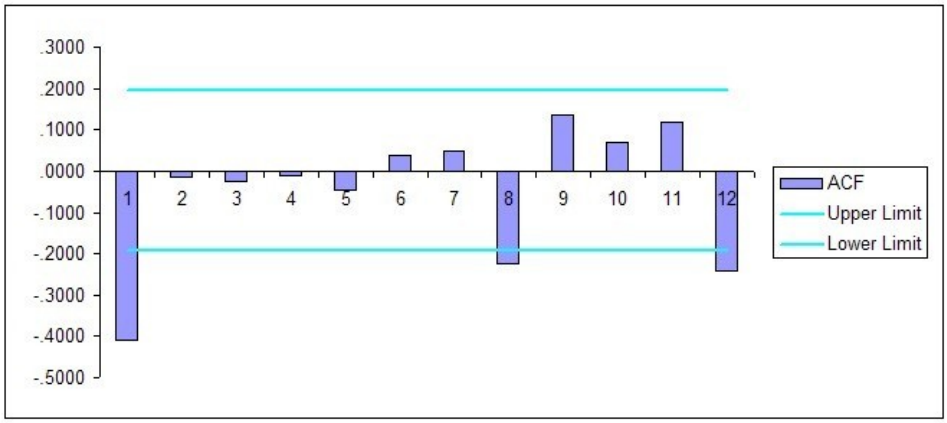
63) A time series whose 24-quarter lag correlogram shows no tendency to diminish towards zero can be said to
A) have a trend term.
B) be nonstationary.
C) have a long memory.
D) be serially correlated.
E) All of the options are correct.

Answer: E
Difficulty: 2 Medium
Topic: Correlograms
Learning Objective: 2-09 Explain how an autocorrelation function (ACF) can be useful in forecasting.
Accessibility: Keyboard Navigation
Gradable: automatic

64)



Autocorrelation Function for GDP

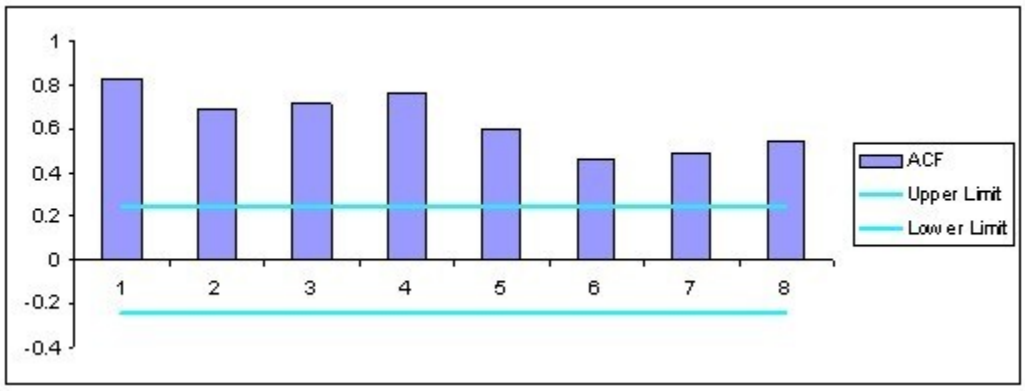


Autocorrelation function for GDP with second differencing

Refer to the autocorrelation functions for Gross Domestic Product (GDP) presented above. The shape of these autocorrelation functions
A) indicates that GDP has a trend.
B) indicates that GDP is seasonal.
C) indicates that GDP is stationary.
D) indicates that GDP is not autocorrelated.

Answer: A
Difficulty: 1 Easy
Topic: Correlograms
Learning Objective: 2-09 Explain how an autocorrelation function (ACF) can be useful in forecasting.
Accessibility: Keyboard Navigation
Gradable: automatic

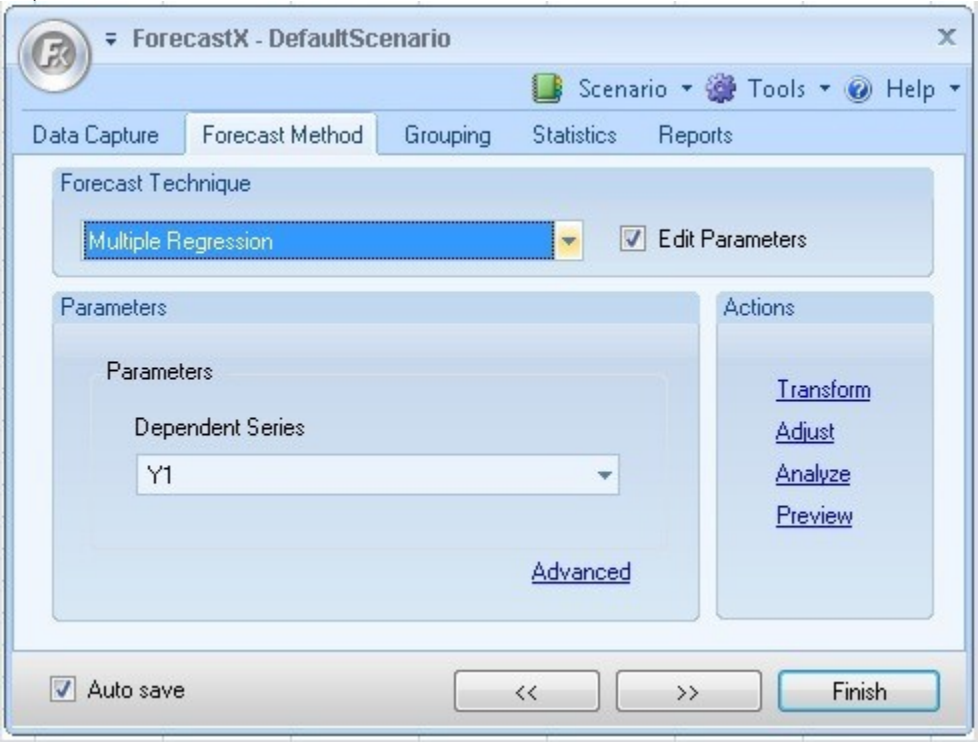
65)



The ACF above represents home sales between March of 1989 and December of 2004. Does the data show evidence of trend?
A) Yes.
B) No.
C) There is too much seasonality to identify if there is a trend or if there is no trend.
D) This data is stationary and the question is not relevant.

Answer: A
Difficulty: 1 Easy
Topic: Total Houses Sold: Exploratory Data Analysis and Model Selection
Learning Objective: 2-09 Explain how an autocorrelation function (ACF) can be useful in forecasting.
Accessibility: Keyboard Navigation
Gradable: automatic

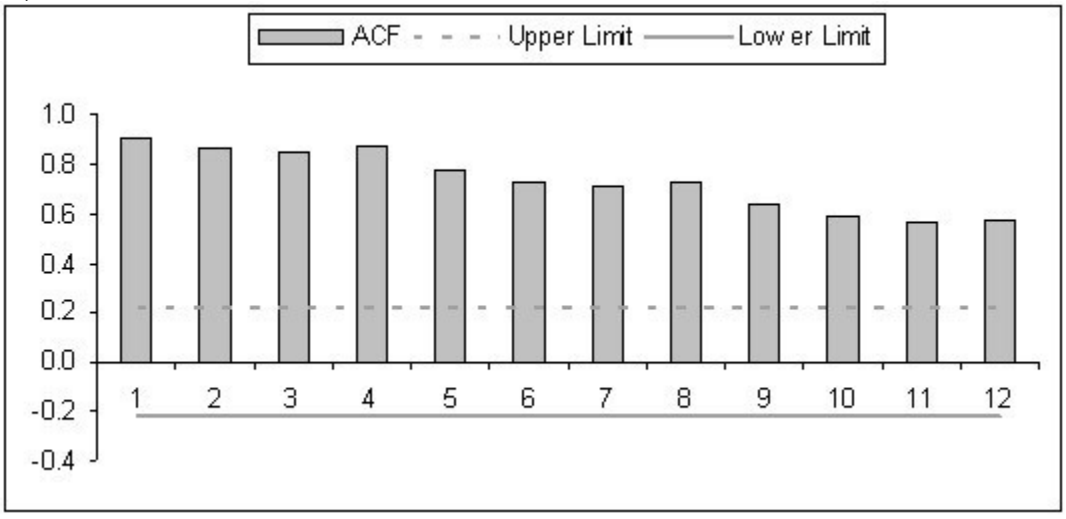
66)



Above is a ForecastX dialog box. How would you begin to check the autocorrelation function?
A) In the "Forecasting Technique" drop-down menu, select autocorrelation.
B) In the "Transform" button, select autocorrelation.
C) Use the "Analyze" button to access the ACF manipulation panel.
D) Use the "Advanced" button to access the lag structure setup panel.
E) None of the options are correct.

Answer: C
Difficulty: 1 Easy
Topic: Using ForecastX to Find Autocorrelation Functions
Learning Objective: 2-09 Explain how an autocorrelation function (ACF) can be useful in forecasting.
Accessibility: Keyboard Navigation
Gradable: automatic

67)



The ACF for The Gap sales is shown above. There is clear evidence in the ACF that
A) the data is stationary.
B) there is a strong trend in the data.
C) the data is too strongly correlated to identify trend.
D) there is no seasonality in the data.
E) Gap sales have fallen in the last 12 periods.

Answer: B
Difficulty: 1 Easy
Topic: Data Analysis of the Gap Sales Data
Learning Objective: 2-09 Explain how an autocorrelation function (ACF) can be useful in forecasting.
Accessibility: Keyboard Navigation
Gradable: automatic

68) A research study has reported that there is a correlation of r = -0.59 between the eye color (brown, green, blue) of an experimental animal and the amount of nicotine that is fatal to the animal when consumed. This indicates
A) nicotine is less harmful to one eye color than the others.
B) the lethal dose of nicotine goes down as the eye color of the animal changes.
C) one must always consider the eye color of animals in making statements about the effect of nicotine consumption.
D) the researchers need to do further study to explain the causes of this negative correlation.
E) the researchers need to take a course in statistics because correlation is not an appropriate measure of association in this situation.

Answer: E
Explanation: Correlation cannot be computed with nominal variables.
Difficulty: 1 Easy
Topic: Correlograms
Learning Objective: 2-09 Explain how an autocorrelation function (ACF) can be useful in forecasting.
Accessibility: Keyboard Navigation
Gradable: automatic

69) If the correlation between body weight and annual income were high and positive, we could conclude that
A) high incomes cause people to eat more food.
B) low incomes cause people to eat less food.
C) high income people tend to spend a greater proportion of their income on food than low income people, on average.
D) high income people tend to be heavier than low income people, on average.
E) high incomes cause people to gain weight.

Answer: D
Difficulty: 1 Easy
Topic: Correlograms
Learning Objective: 2-09 Explain how an autocorrelation function (ACF) can be useful in forecasting.
Accessibility: Keyboard Navigation
Gradable: automatic

70) A study examined the relationship between the sepal length and sepal width for two varieties of an exotic tropical plant. Varieties A and B are represented by x's and o's, respectively, in the following plot. Which of the following statements is FALSE?
A) Considering variety A alone, there is a negative correlation between sepal length and sepal width.
B) Considering variety B alone, the least squares regression line for predicting sepal length from sepal width has a negative slope.
C) Considering both varieties together, there is a positive correlation between sepal length and sepal width.
D) Considering each variety separately, there is a positive correlation between sepal length and sepal width.
E) Considering both varieties together, the least squares regression line for predicting sepal length from sepal width has a positive slope.

Answer: D
Difficulty: 1 Easy
Topic: Correlograms
Learning Objective: 2-09 Explain how an autocorrelation function (ACF) can be useful in forecasting.
Accessibility: Keyboard Navigation
Gradable: automatic

71) The correlation coefficient provides
A) a measure of the extent to which changes in one variable cause changes in another variable.
B) a measure of the strength of the linear association between two categorical variables.
C) a measure of the strength of the association (not necessarily linear) between two categorical variables.
D) a measure of the strength of the linear association between two quantitative variables.
E) a measure of the strength of the linear association between a quantitative variable and a categorical variable.

Answer: D
Difficulty: 1 Easy
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

72) Which of the following correlation coefficients suggests the strongest relationship between the variables?
A) 1.03
B) −.02
C) −.89
D) .75

Answer: C
Difficulty: 1 Easy
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

73) A k-period plot of autocorrelations is called
A) an autocorrelation function.
B) a simple regression line.
C) a product diffusion curve.
D) a scatter plot.

Answer: A
Difficulty: 1 Easy
Topic: Correlograms
Learning Objective: 2-09 Explain how an autocorrelation function (ACF) can be useful in forecasting.
Accessibility: Keyboard Navigation
Gradable: automatic

74) When evaluating a time series of data, it is useful to look at the correlation between successive observations over time. This measure of correlation is called
A) a *t*-distribution
B) an autocorrelation.
C) a paired correlation.
D) a trial correlation.

Answer: B
Difficulty: 1 Easy
Topic: Correlograms
Learning Objective: 2-09 Explain how an autocorrelation function (ACF) can be useful in forecasting.
Accessibility: Keyboard Navigation
Gradable: automatic

75) When we test the significance of a correlation coefficient, the null hypothesis is usually
A) ignored.
B) tested with n-1 degrees of freedom.
C) that the correlation coefficient is equal to zero.
D) that the correlation coefficient is equal to one.

Answer: C
Difficulty: 1 Easy
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

76) The correlation coefficient (r) represents
A) the linear association between two variables.
B) the causal relationship between two variables.
C) the degree of significance in the *t*-test between two variables.
D) the positive association between two variables.

Answer: A
Difficulty: 1 Easy
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

77) The correlation coefficient is also called
A) the variance.
B) Theil's U.
C) the autocorrelation coefficient.
D) the Pearson product-moment correlation coefficient.

Answer: D
Difficulty: 1 Easy
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

78) Correlation coefficients may range in value
A) from zero to four.
B) from zero to 100 percent.
C) from −1 to +1.
D) from zero to one.

Answer: C
Difficulty: 1 Easy
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

79) The correlograms produced by ForecastX™
A) include the autocorrelation plot and the partial autocorrelation plot.
B) are produced only when using cross-sectional data.
C) can be used to run a hypothesis test.
D) should not be used with data sets containing more than 100 observations.

Answer: A
Difficulty: 2 Medium
Topic: Using ForecastX to Find Autocorrelation Functions
Learning Objective: 2-09 Explain how an autocorrelation function (ACF) can be useful in forecasting.
Accessibility: Keyboard Navigation
Gradable: automatic

80) When running a hypothesis test, the process begins by setting up two hypotheses,
A) the average hypothesis and the mean hypothesis.
B) the theoretical hypothesis and the statistical hypothesis.
C) the null hypothesis and the alternative hypothesis.
D) Pearson's hypothesis and the null hypothesis.

Answer: C
Difficulty: 1 Easy
Topic: Hypothesis Testing
Learning Objective: 2-07 Describe three common forms of statistical hypotheses.
Accessibility: Keyboard Navigation
Gradable: automatic

81) The *t*-Distribution (also called the Student's *t*-Distribution)
A) resembles a Chi-Square distribution.
B) resembles a normal distribution.
C) resembles Theil's Distribution.
D) resembles a Gaussian distribution.
E) resembles both a normal and a Gaussian distribution.

Answer: E
Difficulty: 1 Easy
Topic: The Student's t-Distribution
Learning Objective: 2-06 Discuss the normal and Student's t distributions.
Accessibility: Keyboard Navigation
Gradable: automatic

82) The normal distribution
A) is fully defined by three characteristics: the median, the mode, and the mean.
B) is symmetrical around the mean.
C) is identical to the *t*-Distribution.
D) cannot be easily distinguished from a Chi-Square distribution.

Answer: B
Difficulty: 1 Easy
Topic: The Normal Distribution
Learning Objective: 2-06 Discuss the normal and Student's t distributions.
Accessibility: Keyboard Navigation
Gradable: automatic

83) In a normal distribution, what percentage of the area under the curve is included between one standard deviation below the mean and one standard deviation above the mean?
A) 68%
B) 95%
C) 99%
D) There is no standard answer to the question.

Answer: A
Difficulty: 1 Easy
Topic: The Normal Distribution
Learning Objective: 2-06 Discuss the normal and Student's t distributions.
Accessibility: Keyboard Navigation
Gradable: automatic

84) The range of a data set
A) is the difference between the smallest value and the greatest value.
B) is the highest value in the data set.
C) is the difference between zero and the largest value in the data set.
D) is the standard deviation divided by the number of observations.

Answer: A
Difficulty: 1 Easy
Topic: A Statistical Review
Learning Objective: 2-04 Explain the differences between the mean, median, and mode for a set of data.
Accessibility: Keyboard Navigation
Gradable: automatic

85) Using a Student's t-Distribution table and 85 degrees of freedom, what value of t would be associated with finding 97.5 percent of the area in the unshaded region (i.e., the area not in the tail)?
A) zero
B) 99%
C) 1.96
D) 0.050

Answer: C
Difficulty: 1 Easy
Topic: The Student's t-Distribution
Learning Objective: 2-06 Discuss the normal and Student's t distributions.
Accessibility: Keyboard Navigation
Gradable: automatic

86) Consider the following formula:

$$? = \frac{X - \mu_0}{s/\sqrt{n}}$$

This formula represents
A) a calculated Z-statistic.
B) a calculated *t*-statistic.
C) a calculated standard deviation.
D) a calculated correlation coefficient.

Answer: B
Difficulty: 1 Easy
Topic: The Student's t-Distribution
Learning Objective: 2-06 Discuss the normal and Student's t distributions.
Accessibility: Keyboard Navigation
Gradable: automatic

87) Consider the following equation:

$$? = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sqrt{\left[\sum (X - \bar{X})^2\right]\left[\sum (Y - \bar{Y})^2\right]}}$$

This equation represents
A) a calculated *t*-statistic.
B) the calculated standard deviation.
C) the calculated Pearson product-moment correlation coefficient.
D) the calculated Z-statistic.

Answer: C
Difficulty: 1 Easy
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

88) Which of the following is a measure of central tendency?
A) median
B) standard deviation
C) *t*-statistic
D) autocorrelation coefficient

Answer: A
Difficulty: 1 Easy
Topic: A Statistical Review
Learning Objective: 2-04 Explain the differences between the mean, median, and mode for a set of data.
Accessibility: Keyboard Navigation
Gradable: automatic

89) A cyclical pattern
A) is represented by wavelike upward and downward movements of the data around the long-term trend.
B) occurs in a time series when there is a regular variation in the level of the data that repeats itself at the same time each year.
C) is a long-term change in the level of the data.
D) contains the fluctuations that are not part of the other three components.

Answer: A
Difficulty: 1 Easy
Topic: Trend, Seasonal and Cyclical Data Patterns
Learning Objective: 2-02 Distinguish between trend, seasonal, and cyclical data patterns.
Accessibility: Keyboard Navigation
Gradable: automatic

90) A seasonal pattern
A) is represented by wavelike upward and downward movements of the data around the long-term trend.
B) occurs in a time series when there is a regular variation in the level of the data that repeats itself at the same time each year.
C) is a long term change in the level of the data.
D) contains the fluctuations that are not part of the other three components.

Answer: B
Difficulty: 1 Easy
Topic: Trend, Seasonal and Cyclical Data Patterns
Learning Objective: 2-02 Distinguish between trend, seasonal, and cyclical data patterns.
Accessibility: Keyboard Navigation
Gradable: automatic

91) A trend in a time series
A) is represented by wavelike upward and downward movements of the data around the long-term trend.
B) occurs in a time series when there is a regular variation in the level of the data that repeats itself at the same time each year.
C) is a long term change in the level of the data.
D) contains the fluctuations that are not part of the other three components.

Answer: C
Difficulty: 1 Easy
Topic: Trend, Seasonal and Cyclical Data Patterns
Learning Objective: 2-02 Distinguish between trend, seasonal, and cyclical data patterns.
Accessibility: Keyboard Navigation
Gradable: automatic

92) Fit refers to
A) how well the model works in the forecast horizon.
B) how well the model works retrospectively.
C) how accurate the measures of central tendency will be.
D) the accuracy of the seasonal component of a forecasting model.

Answer: B
Difficulty: 1 Easy
Topic: The Forecast Process
Learning Objective: 2-01 Explain a process for developing forecasts.
Accessibility: Keyboard Navigation
Gradable: automatic

93) Accuracy refers to
A) how well the model works in the forecast horizon.
B) how well the model works retrospectively.
C) how accurate the measures of central tendency will be.
D) the accuracy of the seasonal component of a forecasting model.

Answer: A
Difficulty: 1 Easy
Topic: The Forecast Process
Learning Objective: 2-01 Explain a process for developing forecasts.
Accessibility: Keyboard Navigation
Gradable: automatic

94)

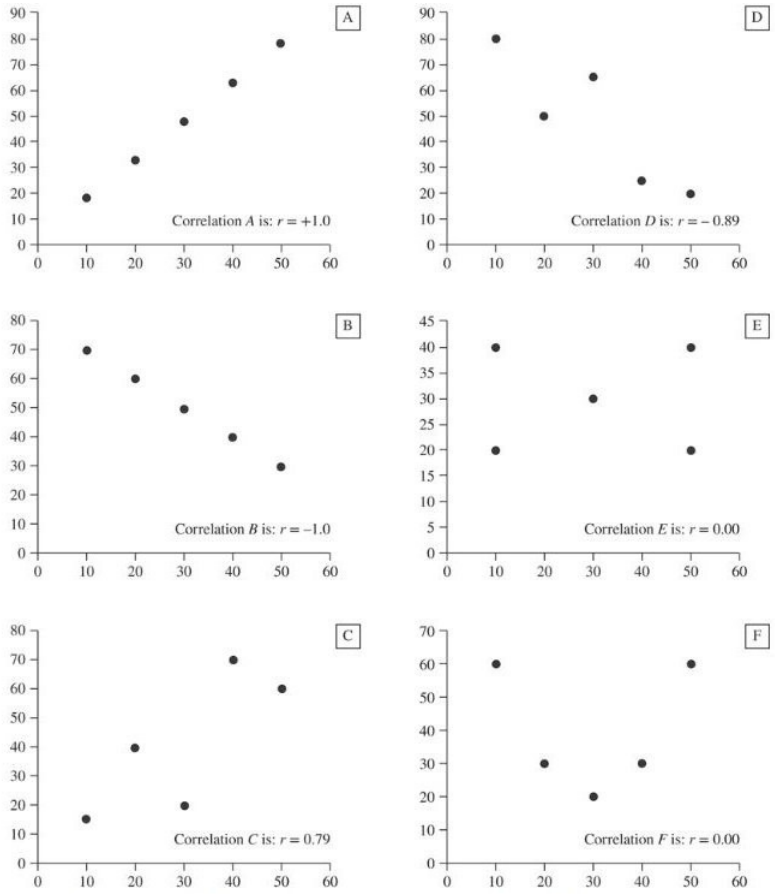


FIGURE 2.6 Representative scatterplots with the corresponding correlation coefficients. These scatterplots show correlation coefficients that range from a perfect positive correlation (*A*) and a perfect negative correlation (*B*) to zero correlations (*E* and *F*).

Which frame of the correlation diagram (A through F) represents a perfect positive linear correlation?
A) A
B) B
C) C
D) D
E) E
F) F

Answer: A
Difficulty: 1 Easy
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

95)

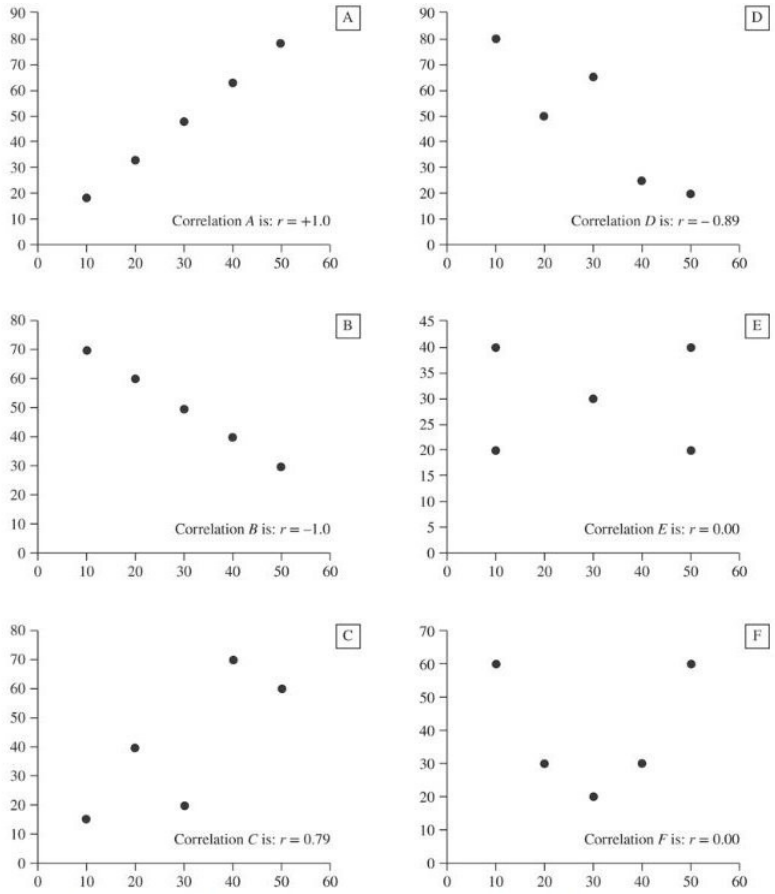


FIGURE 2.6 Representative scatterplots with the corresponding correlation coefficients. These scatterplots show correlation coefficients that range from a perfect positive correlation (*A*) and a perfect negative correlation (*B*) to zero correlations (*E* and *F*).

Which frame of the correlation diagram (A through F) represents a perfect inverse linear correlation?
A) A
B) B
C) C
D) D
E) E
F) F

Answer: B
Difficulty: 1 Easy
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

96)

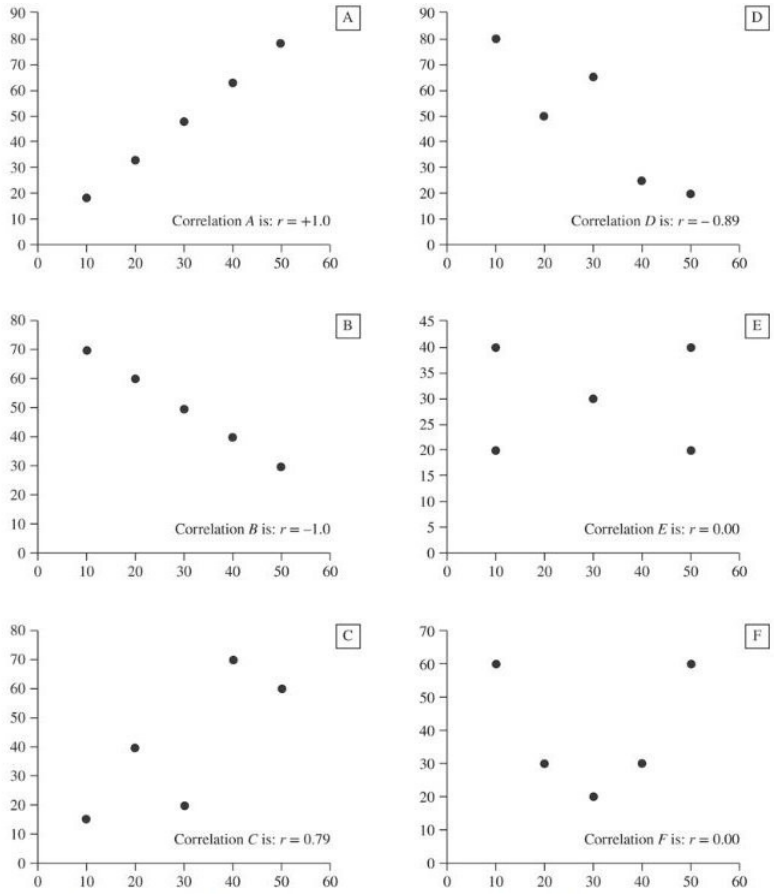


FIGURE 2.6 Representative scatterplots with the corresponding correlation coefficients. These scatterplots show correlation coefficients that range from a perfect positive correlation (*A*) and a perfect negative correlation (*B*) to zero correlations (*E* and *F*).

Which frame of the correlation diagram (A through F) represents an imperfect positive linear correlation?
A) A
B) B
C) C
D) D
E) E
F) F

Answer: C
Difficulty: 1 Easy
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic

97)

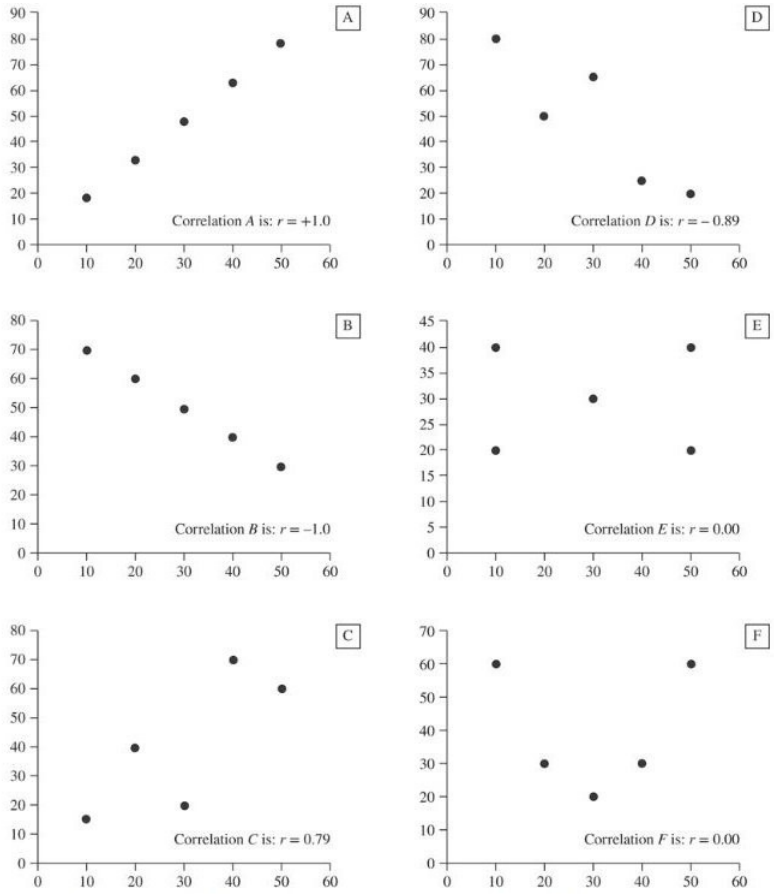


FIGURE 2.6 Representative scatterplots with the corresponding correlation coefficients. These scatterplots show correlation coefficients that range from a perfect positive correlation (*A*) and a perfect negative correlation (*B*) to zero correlations (*E* and *F*).

Which frame of the correlation diagram (A through F) represents an imperfect negative linear correlation?
A) A
B) B
C) C
D) D
E) E
F) F

Answer: D
Difficulty: 1 Easy
Topic: Correlation
Learning Objective: 2-08 Explain what a statistical correlation measures.
Accessibility: Keyboard Navigation
Gradable: automatic