Practice Test #2

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

Suppose you are to form a scatterplot by collecting data for the given pair of variables. Determine the likely direction, form, and scatter.

1) Age, wrinkles
   Answer: Scatterplot: positive, linear, some scatter
   Page Ref: 116-130

2) How tired you are, hours of sleep
   Answer: Scatterplot: negative, linear, some scatter
   Page Ref: 116-130

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Determine whether the scatterplot shows little or no association, a negative association, a linear association, a moderately strong association, or a very strong association (multiple associations are possible).

3) [Diagram of scatterplot]
   A) Little or no association
   B) Positive association, moderately strong association
   C) Positive association, linear association
   D) Negative association, linear association
   E) Negative association, moderately strong association

   Answer: A
   Page Ref: 116-130

Fill in the missing information.

4) \[
\begin{array}{cccc}
  x & s_x & -y & s_y & r & \hat{y} = b_0 + b_1x \\
  ? & 15 & 4 & 9 & 4 & \hat{y} = -20 + 2.4x \\
\end{array}
\]
   A) 3  B) 5  C) 10  D) 9  E) 7

   Answer: C
   Page Ref: 141-155
SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

The graph shows the relationship between the number of games won by a Chicago basketball team and the average salary (in millions) for a ten-year period. Use the scatterplot, the residuals plot, and part of the regression analysis to answer the question.

5) Is a linear model appropriate here?
   Answer: Yes, the relationship is strong and reasonably straight.
   Page Ref: 141-155

6) Interpret the meaning of R² in this context.
   Answer: The salary explains 65% of the variation in wins.
   Page Ref: 141-155

7) What is the correlation between salary and average wins?
   Answer: 0.806
   Page Ref: 141-155
8) What would be a good prediction for the average wins for a team that is 3 standard deviations above average in games won?

Answer: 2.418 SDs above the average.

Page Ref: 141–155

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Predict the slope for the example regression model.

9) If you create a regression model for predicting the weight of a motorcycle (in pounds) from its length (in feet), is the slope most likely to be 0.8, 8, 80, 800, or 8000?

A) 800   B) 80   C) 8000   D) 8   E) 0.8

Answer: B

Page Ref: 141–155

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

Explain the misinterpretation for the given case.

10) A sociology student does a study to determine whether people who exercise live longer. He claims that someone who exercises 7 days a week will live 15 years longer than someone who doesn’t exercise at all.

Answer: Predictions based on a regression line are for average values of $y$ for a given $x$. The actual life expectancy will vary around the prediction.

Page Ref: 141–155

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Answer the question.

11) The scatterplot below displays the yearly production in millions of pounds of flue-cured tobacco in the U.S. For what range of years is a linear model appropriate?

A) A linear model should be used for each pair of adjacent data points.
B) One linear model for 1919 through about 1960 and another linear model for about 1960 through 2000.
C) A single linear model is appropriate for the entire data set.
D) A linear model should not be used for any part of the data.
E) None of these

Answer: B

Page Ref: 166–177
12) Which of the labeled points below are outliers?

A) Points A, B, C, and D
B) Points A and C
C) Point A
D) Points A, C, and D
E) Points C and D

Answer: D
Page Ref: 166-177

13) Which of the labeled points below are influential points?

A) Points A and C
B) Points A and D
C) Points A, C, and D
D) Point D
E) Points A, B, C, and D

Answer: D
Page Ref: 166-177
14) Which of the labeled points below will exert the largest leverage on a linear model of the data?

A) Point C  B) Point A  C) Point D  D) Point E  E) Point B

Answer: D
Page Ref: 166–177

15) Which of the following scatterplots of residuals suggests that a linear model may not be applicable?

A) 

B) 

Page Ref: 166–177
Answer: A
Page Ref: 166-177
16) The figure below examines the association between life expectancy and computer ownership for several countries. Also shown are the equation and $R^2$ value from a linear regression analysis. What is the best conclusion to draw from the figure?

A) Computer ownership promotes health and long life, probably due to the greater access that computer owners have to health information on the world-wide web.

B) Persons who live longer buy more computers over the course of their longer lifetimes.

C) Clearly, there must be some as-yet unknown health benefit associated with using computers.

D) Exposure to the radiation from computer monitors is causing a clear decline in life expectancy.

E) Although the association is strong, computer ownership probably does not promote longevity. Instead, national per capita wealth is probably a lurking variable that drives both life expectancy and computer ownership.

Answer: E
Page Ref: 166-177

17) Criticize the following simulation: A student simulates math test scores by choosing a random number between 0 and 100.

A) The simulation will not model the real situation. It fails to account for the amount of time each student spent studying for the exam.

B) The simulation probably will not model the real situation. For example, the simulation will predict just as many scores between 10 and 20 as between 70 and 80. In reality, the distribution of grades will not be so uniform.

C) The simulation cannot model the real situation. The test performance of an individual student is inherently unpredictable.

D) The simulation probably will not model the real situation. Most students dislike math.

E) The simulation should model the real situation.

Answer: B
Page Ref: 219-225
18) Staff members for a local alderman go to one of the city’s election polling places in his district and interview all the residents they can find.

Answer: Population—Registered voters in the local alderman’s district.
Parameter—Not clear/Percentage of voters favoring various issues?
Sampling Frame—All district residents.
Sample—Cluster sample; one of the city’s election polling places
Method—Convenience sample within each cluster
Bias—Parameter(s) of interest not clear. Sampling within clusters is not random and may bias results.

Page Ref: 231-243

19) At a college there are 120 freshmen, 90 sophomores, 110 juniors, and 80 seniors. A school administrator selects a random sample of 12 of the freshmen, a random sample of 9 of the sophomores, a random sample of 11 of the juniors, and a random sample of 8 of the seniors. She then interviews all the students selected. Identify the type of sampling used in this example.

A) Voluntary response  B) Stratified sampling  C) Attempted census  D) Cluster sampling

Answer: B

Page Ref: 231–243

20) At the local college a survey was being done on whether or not the students liked the cafeteria food. The survey was located in the college newspaper and was to be filled out and sent to the editing office. Identify the type of sampling used in this example.

A) Stratified sampling  B) Attempted census  C) Voluntary response  D) Cluster sampling

Answer: C

Page Ref: 231–243

21) Given that several 18–year olds are still in high school, do you think they are old enough to vote?

Answer: Biased towards no because of "several 18–year olds are still in high school." "Do you think the voting age of 18 is appropriate?"

Page Ref: 231–243

22) Given the university’s great reputation of powerful athletics, do you favor continued funding for athletic scholarships?

Answer: Biased toward yes because of "great reputation". "Do you favor continued funding for the university athletic program?"

Page Ref: 231–243
1) Scatterplot: positive, linear, some scatter
   ID: STATDMID 7.1.2-1
   Diff: 0       Page Ref: 116–130

2) Scatterplot: negative, linear, some scatter
   ID: STATDMID 7.1.2-2
   Diff: 0       Page Ref: 116–130

3) A
   ID: STATDMID 7.1.3-3
   Diff: 0       Page Ref: 116–130

4) C
   ID: STATDMID 8.1.1-1
   Diff: 0       Page Ref: 141–155

5) Yes, the relationship is strong and reasonably straight.
   ID: STATDMID 8.1.3-1
   Diff: 0       Page Ref: 141–155

6) The salary explains 65% of the variation in wins.
   ID: STATDMID 8.1.3-2
   Diff: 0       Page Ref: 141–155

7) 0.806
   ID: STATDMID 8.1.3-3
   Diff: 0       Page Ref: 141–155

8) 2.418 SDs above the average.
   ID: STATDMID 8.1.3-4
   Diff: 0       Page Ref: 141–155

9) B
   ID: STATDMID 8.1.4-1
   Diff: 0       Page Ref: 141–155

10) Predictions based on a regression line are for average values of y for a given x. The actual life expectancy will vary around the prediction.
    ID: STATDMID 8.1.5-2
    Diff: 0       Page Ref: 141–155

11) B
    ID: STATDMID 9.1.1-2
    Diff: 0       Page Ref: 166–177

12) D
    ID: STATDMID 9.1.3-1
    Diff: 0       Page Ref: 166–177

13) D
    ID: STATDMID 9.1.3-3
    Diff: 0       Page Ref: 166–177

14) D
    ID: STATDMID 9.1.3-5
    Diff: 0       Page Ref: 166–177

15) A
    ID: STATDMID 9.1.5-2
    Diff: 0       Page Ref: 166–177
16) E
   ID: STATDM1D 9.1.4-1
   Diff: 0       Page Ref: 166-177

17) B
   ID: STATDM1D 11.1.1-4
   Diff: 0       Page Ref: 219-225

18) Population-Registered voters in the local alderman's district.
   Parameter-Not clear/Percentage of voters favoring various issues?
   Sampling Frame-All district residents.
   Sample-Cluster sample; one of the city's election polling places
   Method-Convenience sample within each cluster
   Bias-Parameter(s) of interest not clear. Sampling within clusters is not random and may bias results.
   ID: STATDM1D 12.1.1-3
   Diff: 0       Page Ref: 231-243

19) B
   ID: STATDM1D 12.1.2-2
   Diff: 0       Page Ref: 231-243

20) C
   ID: STATDM1D 12.1.2-3
   Diff: 0       Page Ref: 231-243

21) Biased towards no because of "several 18-year olds are still in high school." "Do you think the voting age of 18 is appropriate?"
   ID: STATDM1D 12.1.3-3
   Diff: 0       Page Ref: 231-243

22) Biased toward yes because of "great reputation". "Do you favor continued funding for the university athletic program?"
   ID: STATDM1D 12.1.3-4
   Diff: 0       Page Ref: 231-243