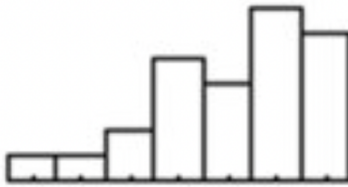# Quiz 1 for Stat 11

Solutions

2.17.23

Name: _____

Part 1: Multiple Choice

X. Which is TRUE of the data shown in the histogram below?

I. The distribution is skewed to the right. II. The mean is probably smaller than the median. III. We should use median and IQR to summarize these data.

  (a) I only

  (b) II only

  (c) III only

(d) II and III only

  (e) I, II, and III

X. The following box plots show the closing share prices for a sample of technology companies on the first trading days in August 2007 and in August 2002.

Boxplot of Aug. 2007, Aug. 2002
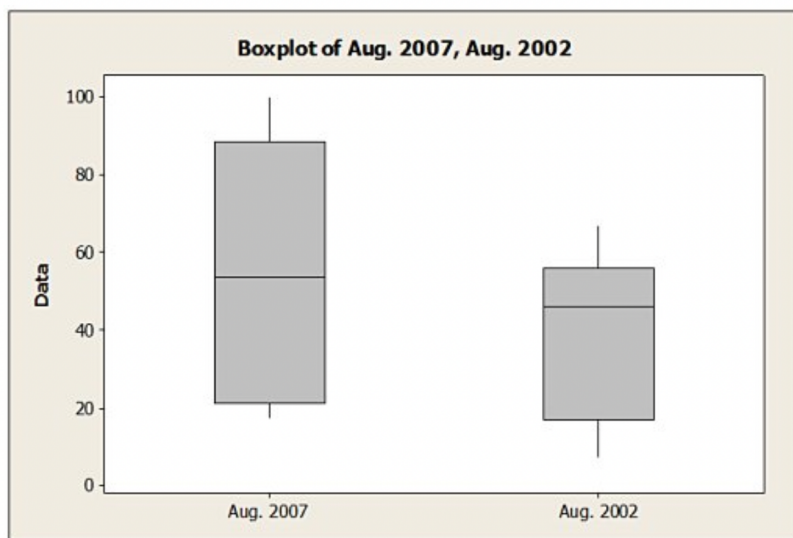
Which of the following statements is TRUE?

(a) The median closing share price is higher in August 2007 compared to August 2002.

(b) Closing prices are more variable in August 2007 compared to August 2002.

(c) The distribution of closing prices in August 2007 appears more symmetric than the distribution of closing prices in August 2002.

(d) Both A and B

(e) All of the above

X. All but one of the statements below contain a mistake. Which one could be TRUE?

(a) The correlation between your golf score and the number of hours you practice is 0.36.

(b) The number of apricots on a tree and the amount of fertilizer have a 1.12 correlation.

(c) There is a strong correlation between type of preferred pet and income level.

(d) The correlation between the height of a bean plant and the day is 0.78 in/day.

X. A company's sales increase by the same amount each year. This growth is _____

(a) exponential

(b) logarithmic

(c) power

(d) linear

(e) quadratic

X. A regression analysis of company profits and the amount of money the company spent on advertising found $r^2 = 0.72$. Which of these is TRUE?

I. This model can correctly predict the profit for 72% of companies.

II. On average, about 72% of a company's profit results from advertising.

III. On average, companies spend about 72% of their profits on advertising.

(a) none

(b) I only

(c) II only

(d) III only

(e) I and III

X. The residuals plot for a linear model is shown. Which is TRUE?

(a) The linear model is no good since some residuals are large.

(b) The linear model is okay because the association between the two variables is fairly strong.

(c) The linear model is no good since the correlation is near 0.

(d) The linear model is okay because about the same number of points are above and below the it.

(e) The linear model is no good because of the curve in the residuals.

X. A residual plot that has no pattern is a sign of what?

(a) The original data is curved and the regression line is a good model.

(b) The original data is curved and the regression line is not a good model.

(c) The model is not a good one, because there is no pattern.

(d) The original data is straight and the regression line is not a good model.

(e) The original data is straight and the regression line is a good model.

X. Which of the following is NOT a goal of re-expressing data?

(a) Make the distribution of a variable more symmetric.

(b) Make the form of a scatter plot more nearly linear.

(c) Make the scatter in a scatter plot spread out evenly rather than following a fan shape.

(d) Make the spread of several groups more alike.

(e) All of these are goals of re-expressing data.

X. A clothing store uses comment cards to get feedback from its customers about newly added items. It recently introduced plus size fashion wear. Customers who purchased the items were asked to fill out an online comment survey giving 10% off the next purchase. The data are summarized in the table below. What percentage of customers were at least satisfied with the item(s) purchased (Satisfied or Very satisfied)?

Response	Frequency
Very satisfied	15
Satisfied	30
Less than fully satisfied	12
Not satisfied	4

(a) 68.9%

(b) 73.8%

(c) 49.2%

(d) 26.2%

(e) 24.5%

X. A company's manufacturing process uses 500 gallons of water at a time. A "scrubbing" machine then removes most of a chemical pollutant before pumping the water into a nearby lake. To meet federal regulations the treated water must not contain more than 80 parts per million (ppm) of the chemical. Because a fine is charged if regulations are not met, the company sets the machine to attain an average of 75 ppm in the treated water. The machine's output can be described by a normal model with standard deviation 4.2 ppm.

What percent of the batches of water discharged exceed the 80 ppm standard?

(a) 3.89%

(b) 11.7%

(c) 8.83%

(d) 1.17%

(e) 88.3%

Part 2: Fill in the blank

X. A regional survey was carried out to gauge public opinion on the controversial Arizona Immigration Law

(results shown below). _____ percent of the public oppose the law.

| Response	Democrat	Republican	Independent
Favor	50	93	35
Oppose	85	45	60
Don't know	5	7	20

A: 44.5%

X. Based on data collected from its production processes, Crosstiles Inc. determines that the breaking strength of its most popular porcelain tile is normally distributed with a mean of 400 pounds per square inch and a standard deviation of 12.5 pounds per square inch. Based on the 68-95-99.7 Rule, about

_____ percent of the popular porcelain tile will have breaking strengths between 375 and 425 pounds per square inch.

A: 97.35%

X. The box plots show prices of used cars (in thousands of dollars) advertised for sale at three different car dealers.
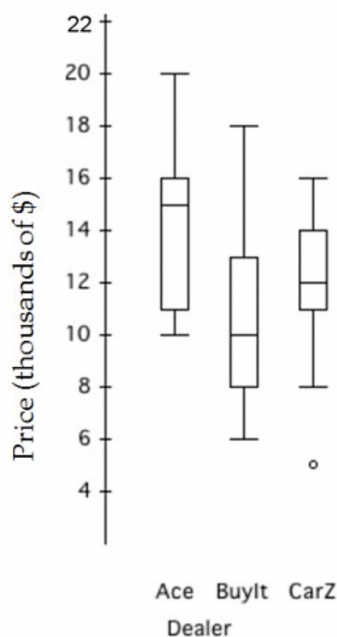
_____ is the dealer with the smallest price range of approximately _____ dollars.

A: Ace, 10,000

X. At a local manufacturing plant, employees must complete new machine set ups within 30 minutes. Suppose new machine set-up times can be described by a normal model with a mean of 22 minutes and a standard deviation of 4 minutes.

_____ percent of new machine set ups take more than 30 minutes.

A: 2.5%

X. Suppose the time it takes to process phone orders in a small florist/gift shop is normally distributed with a

mean of 6 minutes and a standard deviation of 1.24 minutes. An order that processes in _____ minutes separates the 5% of orders that take the most time to process from the rest of the orders.

A: 7.8

Part 3: Free response

A small independent organic food store offers a variety of specialty coffees. To determine whether price has an impact on sales, the managers kept track of how many pounds of each variety of coffee were sold last month.

	PRICE PER POUND	POUNDS SOLD
	$ 3.99	75
	$ 5.99	60
	$ 7.00	65
	$ 12.00	45
	$ 4.50	80
	$ 7.50	70
	$ 15.00	25
	$ 10.00	35
	$ 12.50	40
	$ 8.99	50
Mean	$ 8.75	54.50
Standard Deviation	$ 3.63	18.33
Correlation	− 0.927	

X. Interpret the correlation within the context of the data.

_____

_____

_____

A: strong negative linear, association price increases, lbs sold decreases, provided association is actually linear (verify with plot)

X. Most likely, in a simple linear regression model for this data, _____ is the predictor variable and

_____ is the response variable.

A: price, lbs sold

Data were collected on monthly sales revenues (in $1,000s) and monthly advertising expenditures ($100s) for a sample of drug stores. The regression line relating revenues (Y) to advertising expenditure (X) is estimated to be $\hat{y}_i = -48.3 + 9.00x_i$.

X. Interpret the slope within the context of this data. (Make sure you specify the correct units of measurement.)

A: for every \$100 spent on advertising, the expected/ave/predicted sales revenue increases by \$9000