First Major - 053

Question.1. (6-Points)
Answer True or False

Solutions

True 1. Recently, Al-Riyadh bank manager pulled a sample of customer accounts and recorded data for the total number of transactions during the previous 30 days. The data collected would be considered time series data.---True

True 2. For the same data, a graph of a relative frequency distribution will look exactly the same as a graph of the frequency distribution:---True

False 3. The data collected by mail questionnaire is a secondary data, since it is collected by sending emails to the selected persons using a computer:---False

False 4. A distribution is said to be symmetric when the sample mean and the population mean are equal.---False

True 5. A tire store manager has collected data showing the number of tires of each brand sold during the past month. A pie chart might be effective in graphically illustrating which brands tend to sell best at this store.---True

True 6. One of the reasons that the standard deviation is preferred as a measure of variation over the variance is that the standard deviation is measured in the original units.:---True

Question.2. (6-Points)
Answer the following questions by choosing the right answer.

1. When a manager of a certain market prepares a series of charts and graphs to all people who invested in that market during the past five years, he is using which general category and statistical analysis?

   a. Quantitative statistics
   b. Inferential statistics
   c. Descriptive statistics
   d. None of the above

2. A study of middle to upper level mangers is undertaken to investigate the relationship between salary level and years of work experience. An appropriate graph to display the relationship between the two variables is:

   a. Histogram
   b. Scatter diagram
   c. Stem – and –leaf diagram
   d. Bar chart

✳ Solutions ✳

3. The advantage of using the interquartile range versus the range as a measure of variation is:

   a. It is easier to compute.
   b. It utilizes all the data in its computation.
   c. It gives a value that is closer to the true variation.
   ➤ **(d.)** It is less affected by extremes in the data.

4. If a distribution for a quantitative variable is thought to be nearly symmetric with very little variation:

   a. The width of the box in a box and whisker plot will be quite wide but the whisker will be very short.
   ➤ **(b.)** The left and right-hand box on a box and whisker plot will be approximately equal distance from the median.
   c. The whiskers on a box and whisker plot should be about half as long as the box is wide.
   d. None of the above

5. An effective graph for displaying the growth trend in sales for a company is:

   a. A histogram.
   b. An ogive.
   ➤ **(c.)** A line chart.
   d. A stem & leaf diagram.

6. Which of the following is not considered a statistical sampling method?

   a. Cluster sampling
   b. Systematic sampling
   c. Stratified sampling
   ➤ **(d.)** Convenience sampling

✶ Solutions✶

Question3. (6 +6 = 12–Points)

a. When designing a telephone survey, one of the steps is to formulate the survey questions, explain how to take care of this step?

We have to :

1. make questions clear and unambiguous   ②pts each

2. Use universally – accepted definitions

3. Limit the number of questions.

b. One of the data collection methods is "Direct observation and personal Interviews". Write down three disadvantages of this method.

                 ②pts each·

1. It increases the cost.

2. It needs more time than other methods

3. It depends on the person who makes the interview.

✳ Solutions✳

Question 4(4+2 = 6-Points)
The following observations are the sales for a big mall in 100,000 of SR for the previous 20
months:

| 1.91 | 1.92 | 1.93 | 1.95 | 2.00 | 2.00 | 2.01 | 2.05 | 2.07 | 2.09 |
| 2.09 | 2.10 | 2.12 | 2.15 | 2.16 | 2.24 | 2.27 | 2.28 | 2.30 | 2.33 |

a. construct a stem and leaf diagram for the data

Stem	Leaves
19	1 2 3 5
20	0 0 1 5 7 9 9 ④pts
21	0 2 5 6
22	4 7 8
23	0 3

b. Based on the stem-and-leaf diagram , write a comment about the shape of the dada
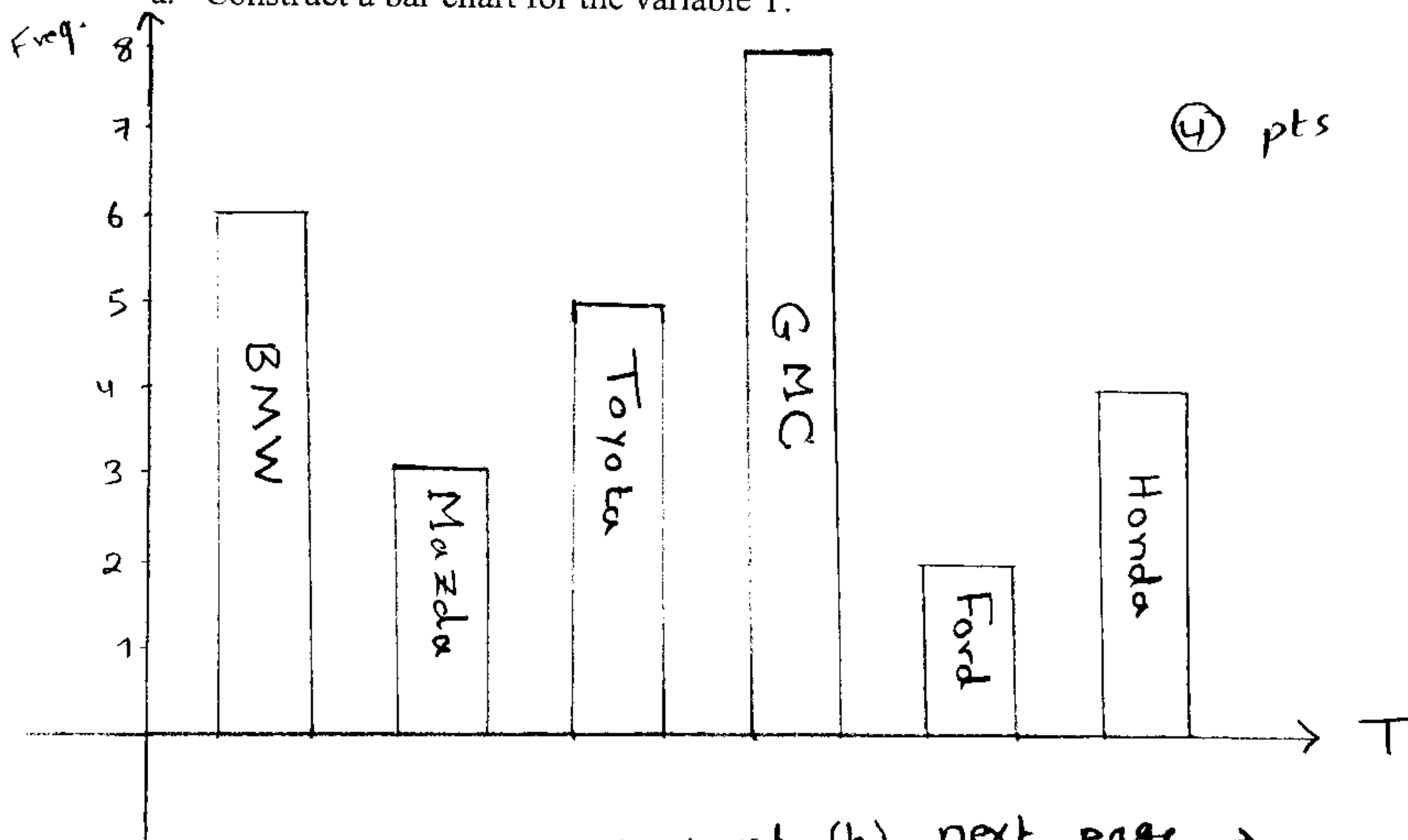distribution

Comment. It is right skewed   ② pts

Question 5(4+2 = 6-Points)
The following table represents the type of cars (T) and the number of sold cars (N) of each type in
one of the big agencies

T	BMW	Mazda	Toyota	GMC	Ford	Honda
N	6	3	5	8	2	4

a. Construct a bar chart for the variable T.

④ pts

Look ⟹ part (b) next page ⟹

b. The mode = GMC   ② pts

✱ Solutions ✱

b. Find the appropriate measure of central tendency for the variable N.   *The mode = GMC*
②pts

Question 6(24-Points)

The following sample data reflect electricity bills (in SR) for ten households in Al-Khubar in March:

118.2   67.88   88.42   110.34   133.4   78.9   144.56   127.89   89.34   129.1

Given that: $\sum x = 1088.03$, $\sum x^2 = 124494.6077$

Find the following

1. Mean

$$\bar{x} = \frac{\sum x_i}{n} = \frac{1088.03}{10} = 108.803 \quad ② \text{ pts}$$

2. Median   The sorted data: $67.88, 78.9, 88.42, 89.34, 110.34$
$118.2, 127.89, 129.1, 133.4, 144.56$

3. Mode $n = 10 \rightarrow$ even $\Rightarrow Md = \frac{110.34 + 118.2}{2} = \frac{228.54}{2} = 114.27 \quad ②\text{pts}$

No mode exist $\}①\text{ pt}$

4. Standard deviation (Hint: $S = \sqrt{\dfrac{\sum(x-\bar{x})^2}{n-1}} = \sqrt{\dfrac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}} = \sqrt{\dfrac{\sum x^2 - n(\bar{x})^2}{n-1}}$)

$$S = \sqrt{\frac{124494.6077 - (10)(108.803)^2}{10-1}}$$

$$= \sqrt{\frac{61136.67961}{9}}$$

$$= \sqrt{6792.9977}$$

$$= 26.0633 \qquad \} ① \text{ pt}$$

③ pts

5. The interquartile range

$$IQR = Q_3 - Q_1$$

②pts $\left\{\begin{array}{l} Q_3 = P_{.75} \Rightarrow p = 75 \quad \therefore i = \frac{75}{100}(10+1) = 8.25 \\ \\ \Rightarrow Q_3 = x_8 + .25(x_9 - x_8) = 129.1 + .25(133.4 - 129.1) = 130.175 \end{array}\right.$

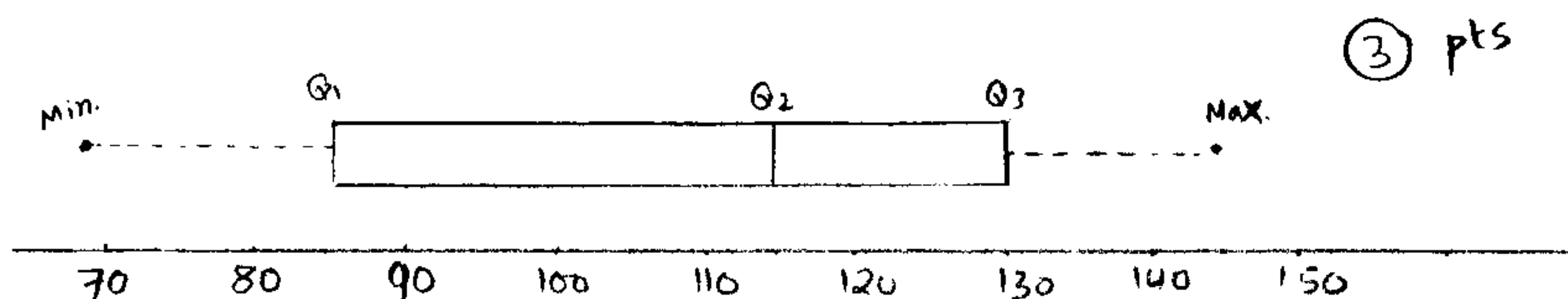②pts $\left\{\begin{array}{l} Q_1 = P_{.25} \Rightarrow p = 25 \quad \therefore i = \frac{25}{100}(10+1) = 2.75 \\ \\ Q_1 = x_2 + .75(x_3 - x_2) = 78.9 + .75(88.42 - 78.9) = 86.04 \end{array}\right.$

$$\therefore IQR = 130.175 - 86.04 = 44.135 \quad \} ① \text{ pt}$$

✳ Solutions ✳

6. Develop a box and whisker plot, and <u>comment</u> on the shape.

Min. $= 67.88$, $Q_1 = 86.04$, $Q_2 = Md = 114.27$, $Q_3 = 130.175$, Max $= 144.56$

③ pts

<u>Comment:</u> It is left skewed. $\}$ ① pt

7. Calculate the 90th percentile.

$P-90 \Rightarrow P = 90 \Rightarrow i = \dfrac{90}{100}(10+1) = 9.9$

$$\text{The } 90^{th} \text{ percentile} = X_9 + .9(X_{10} - X_9)$$
$$= 133.4 + .9(144.56 - 133.4)$$
$$= 133.4 + 10.044$$
$$= 143.444.$$

② pts

8. Determine the coefficient of variation for the set and interpret what it measures.

$$CV = \left(\frac{S}{\bar{X}}\right) \cdot 100\% = \left(\frac{26.0633}{108.803}\right) \cdot 100\%$$
$$= 23.95\%$$

$\}$ ① pt

<u>Interpretation:</u> It measures the relative variation of the data to the mean $\}$ ① pt

9. What is the standardized value for the median electricity bills?

$$Z = \frac{X - \bar{X}}{S} = \frac{Md - \bar{X}}{S} = \frac{114.27 - 108.803}{26.0633}$$
$$= 0.2098$$

② pts

$$\simeq 0.21$$