By Anan Elayan

STAT 2311_1212 TEST 2C

NAME___Key 1___ Number _____ Section _____

QUESTION #1: A researcher has obtained the number of hours worked per week during the summer for a sample of 15 students.

	40	25	35	30	20	40	30	20	40	10	30	20	10	5	20
Sort Data	5	10	10	20	20	20	20	(25)	30	30	30	35	40	40	40

Using this data set, compute the following: **Use SD mode**

1. The median   position: $\frac{15+1}{2} = 8$

Median = 25

2. The mean

25

3. The mode

20

4. The range.

$40 - 5 = 35$

5. The sample standard deviation.

11.34

6. The coefficient of variation.

$C.V = \frac{11.34}{25} \times 100\% = 45.36\%$

7. The limits of the fence of the box plot. Is there any outliers? Explain

$Q_1 : 20$, $Q_3 = 35$, $IQR = 15$, $1.5\ IQR = 22.5$

Lower Limit: $20 - 22.5 = 2.5$

Upper : $35 + 22.5 = 57.5$

No value > 57.5 or less than -2.5

$\Rightarrow$ No outlier

1|S T A T 2 3 1 1 _ 1 2 1 2 _ T E S T 2 C
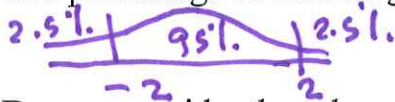
QUESTION #2: Suppose annual salaries for a MM company have a bell-shaped distribution with a mean of $32,500 and a standard deviation of $2,500.

1. The z-score for a sales associate from this store who earns $37,500 is

$$z_{37,500} = \frac{37500 - 32500}{2500} = 2$$

2. The percentage of salaries greater than $37500 is

2.5%

2.5% — 95% — 2.5%
-2 2

3. Do you consider the salary $24000 as a low salary? Why?

$$z_{24000} = \frac{24000 - 32500}{2500} = -3.4 < -3$$

$\Rightarrow$ outlier $\Rightarrow$ Low salary

QUESTION #3: Circle the correct answer

1. Generally, which of the following is the **least appropriate** measure of central tendency for a data set that contains outliers?
   - (a) The mean.
   - b. The median
   - c. The mode
2. An important **measure of location** for categorical data is the
   - a. The mean.
   - b. The median
   - (c) The mode
3. The empirical rule states that, for data having a bell-shaped distribution, the percentage of data values being within two standard deviation of the mean is approximately
   - a. 99.97%
   - (b) 95%
   - c. 68%
   - d. 34%
4. A numerical measure computed from a sample, such as sample mean, is known as a
   - a. Population parameter
   - b. Sample parameter
   - (c) Sample statistic
   - d. Population statistic
5. In computing the pth percentile, if the index i is an integer the pth percentile is the
   - a. Data value in position i
   - b. Data value in position $i+1$
   - (c) Average of data values in positions i and $i + 1$
   - d. Average of data values in positions i and $i - 1$
6. If index i (which is used to determine the location of the pth percentile) is not an integer, its value should be
   - a. squared
   - b. divided by $(n - 1)$
   - c. rounded down
   - (d) rounded up

STAT 2311_1212 TEST 2D

NAME_____*Key 2*_____Number _____Section _____

QUESTION #1: A researcher has obtained the number of hours worked per week during the summer for a sample of 15 students.

	40	25	35	30	20	45	30	20	45	10	30	20	10	45	24
Sort Data	10	10	20	20	20	24	25	(30)	30	30	35	40	45	45	45

Using this data set, compute the following: **Use SD mode**

1. The median

   30

2. The mean

   $\bar{X} = 28.6$

3. The mode

   20, 30, 45

4. The range.

   $45 - 10 = 35$

5. The sample standard deviation.

   $S = 11.75$

6. The coefficient of variation.

   $C.V = \dfrac{11.75}{28.6} \times 100\% = 41.08\%$

7. The limits of the fence of the box plot. Is there any outliers? Explain

   $Q_1 = 20, \quad Q_3 = 40, \quad IQR = 20, \quad 1.5\ IQR = 30$

   Lower Limit: $20 - 30 = -10$

   Upper Limit $40 + 30 = 70$

   No outliers (No values < -10 or > 70)

1|STAT 2311_1212_TEST 2D
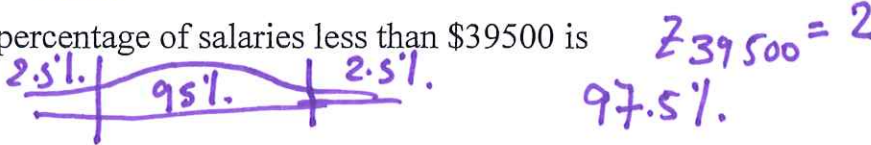
QUESTION #2: Suppose annual salaries for a MM company have a bell-shaped distribution with a mean of $34,500 and a standard deviation of $2,500.

1. The *z*-score for a sales associate from this store who earns $32,000 is

$$Z_{32\,000} = -1$$

2. The percentage of salaries less than $39500 is $Z_{39\,500} = 2$

2.5% | 95% | 2.5%

97.5%

3. Do you consider the salary $42500 as a high salary? Why?

$$Z_{42\,500} = 3.2 > 3 \Rightarrow \text{outlier} \Rightarrow \text{Hish salary}$$

QUESTION #3: Circle the correct answer

1. The empirical rule states that, for data having a bell-shaped distribution, the percentage of data values being within one standard deviation of the mean is approximately
   a. 99.97%
   b. 95%
   c. 68%
   d. 34%

2. Generally, which of the following is the **least appropriate** measure of central tendency for a data set that contains outliers?
   a. The mean.
   b. The median
   c. The mode

3. An important **measure of location** for categorical data is the
   a. The mean.
   b. The median
   c. The mode

4. If index *i* (which is used to determine the location of the *p*th percentile) is not an integer, its value should be
   a. squared
   b. divided by (n − 1)
   c. rounded up
   d. rounded down

5. A numerical measure computed from a population, such as population mean, is known as a
   a. Population parameter
   b. Sample parameter
   c. Sample statistic
   d. Population statistic

6. In computing the *p*th percentile, if the index *i* is an integer the *p*th percentile is the
   a. Data value in position *i*
   b. Data value in position *i+1*
   c. Average of data values in positions *i* and *i − 1*
   d. Average of data values in positions *i* and *i + 1*

2|STAT 2311_1212_TEST 2D