# Class Exercise 4 - Solution

1. **Multiple Choice**

   For each question below, select **all** the statements that are **correct**. Each question has **at least one correct answer, but not necessarily all options are correct**. You will receive **full credit** if and only if you select all correct answers and **no incorrect answers**. Selecting an incorrect option or missing a correct option may result in **partial credit or no credit**.

   A. Which of the following statements about the **Empirical Rule** are true?

      A. A distribution of exam scores is approximately **normal** with mean

      $$\mu = 70 \quad \text{and} \quad \sigma = 8.$$

      Which of the following statements are correct according to the **Empirical Rule**?

      a. ~~Approximately 68% of students scored between 60 and 80.~~
      b. Approximately 99.7% of students scored between 46 and 94.
      c. Approximately 32% of students scored outside the interval from 62 to 78.
      d. Approximately 2.5% of students scored above 86.

   B. Which of the following statements about the **interquartile range (IQR)**, the **coefficient of quartile variation**, and the **Empirical Rule** are true?

      a. The interquartile range (IQR) measures the spread of the middle 50% of the data.
      b. A smaller coefficient of quartile variation indicates less relative variability in the middle half of the data.
      c. ~~The IQR is sensitive to extreme outliers.~~
      d. The coefficient of quartile variation is a unitless measure of variability.

   C. Which of the following statements about choosing appropriate measures of center and spread are true?

      a. If a distribution is strongly right-skewed, the median is generally a better measure of center than the mean.
      b. ~~If a distribution contains a significant outlier, the standard deviation is usually preferred over the interquartile range (IQR) to describe spread.~~
      c. If a distribution contains extreme values, the interquartile range (IQR) is typically a better measure of spread than the standard deviation.
      d. If a distribution is approximately symmetric and unimodal with no outliers, the mean and standard deviation are appropriate summaries.

2. **Leaf Moisture Content**

Due to climate change, chemical changes in leaves are causing moisture levels in some tree species to decrease, making them more susceptible to burning. A group of trees was studied, and the average leaf moisture content (measured as a percentage) was recorded.

| **Leaf Moisture Content (%)** | **Number of Trees** $f_i$ | $m_i$ | $f_i m_i$ | $f_i m_i^2$ | $LTCF$ |
|---|---|---|---|---|---|
| $[0, 10)$ | 45 | 5 | 225 | 1125 | 45 |
| $[10, 20)$ | 47 | 15 | 705 | 10575 | 92 |
| $[20, 30)$ | 33 | 25 | 825 | 20625 | 125 |
| $[30, 40)$ | 25 | 35 | 875 | 30625 | 150 |
| $[40, 50)$ | 10 | 45 | 450 | 20250 | 160 |
| | **160** | | 3080 | 83200 | |

a. Calculate the mean leaf moisture content for this group of trees.

**Solution**

$$\overline{x} = \frac{\sum f_i m_i}{\sum f_i} = \frac{3080}{160} = 19.25\%$$

b. Calculate the sample variance and sample standard deviation of the leaf moisture content. Include units in your answers.

**Solution**

$$s^2 = \frac{1}{n-1}\left[\sum (f_i m_i^2) - \frac{(\sum f_i m_i)^2}{n}\right] = \frac{1}{160-1}\left[83200 - \frac{(3080)^2}{160}\right] = 150.3773\,\%^2$$

$$s = \sqrt{s^2} = \sqrt{150.3773} = 12.2628\%$$

c. Using the Empirical Rule, determine the interval within which you would expect approximately 68% of the data to fall.

**Solution**

$$\overline{x} \pm s$$
$$19.25 \pm 12.2628$$
$$(6.9872, 31.5128)$$

Approximately 68% of the leaves will have a moisture content between 6.9872% and 31.5128%

d. Compute IQR and the CQV for the data above, and interpret it in the context of the problem.

**Solution**

$$Q_1 = \frac{(N+1)P_i}{100} = \frac{(160+1)25}{100} = 40.25 \quad \Rightarrow \quad Q_1 = 5$$
$$Q_3 = \frac{(N+1)P_i}{100} = \frac{(160+1)75}{100} = 120.25 \quad \Rightarrow \quad Q_3 = 25$$

$$\begin{aligned} IQR &= Q_3 - Q_1 \\ &= 25 - 5 \\ &= 20 \end{aligned}$$

$$CQV = \frac{IQR}{Q_1 + Q_3} = \frac{20}{5+25} = 0.6667$$

**Interpretation:** The middle 50% of the leaf moisture content values are spread across approximately 20 percentage points. The coefficient of quartile variation (CQV) of approximately 0.6667 indicates a high level of relative variability in moisture content within this middle portion of the data.