DP Analysis & Approaches (SL/HL)

Maths with Lemon

Lesson 4.3 – Measures of Dispersion

1. Central Tendency and Spread

The mean, median and mode describe the centre of a distribution. The range, interquartile range (IQR), variance and standard deviation describe spread.

$$\bar{x} = \frac{\sum x_i}{n}, \qquad \text{range} = \max(x) - \min(x), \qquad \text{IQR} = Q_3 - Q_1.$$

For frequency data, $\bar{x} = \frac{\sum fx}{\sum f}$. For grouped data use class midpoints to estimate the mean.

2. Quartiles and Outliers

Order the observations. The median divides the data in half; Q_1 and Q_3 describe the lower and upper quarters. Quartile conventions can differ between calculators.

$$L = Q_1 - 1.5\,\text{IQR}, \qquad U = Q_3 + 1.5\,\text{IQR}.$$

Values below L or above U are outliers. On a modified box plot, whiskers end at the most extreme non-outliers; show outliers separately.

Example. For $2, 3, 4, 5, 6, 7, 20$, the median is 5, $Q_1 = 3$, $Q_3 = 7$, $\text{IQR} = 4$. The upper fence is 13, so 20 is an outlier.

3. Variance and Standard Deviation

Deviation from the mean is $x_i - \bar{x}$. Deviations sum to zero, so square them to measure spread.

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n}, \qquad \sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}}.$$

The lesson uses division by n. Some calculators also report sample standard deviation s, which uses $n - 1$; use the correct setting.

Worked example: $2, 3, 5, 9, 11$ has mean 6. Deviations: $-4, -3, -1, 3, 5$; squared deviations: $16, 9, 1, 9, 25$. Thus

$$\sigma^2 = \frac{60}{5} = 12, \qquad \boxed{\sigma = \sqrt{12} \approx 3.46}.$$

4. Frequency Tables

For discrete scores x with frequency f, use

$$\bar{x} = \frac{\sum fx}{N}, \qquad \sigma^2 = \frac{\sum fx^2}{N} - \bar{x}^2, \quad N = \sum f.$$

For grouped intervals, substitute class midpoints and state that the results are estimates.

5. Transformations

If each observation is changed to $y = ax + b$, then

$$\bar{y} = a\bar{x} + b, \qquad \sigma_y = |a|\sigma_x, \qquad \mathrm{Var}(y) = a^2\mathrm{Var}(x).$$

Adding a constant changes the centre but not the spread. Multiplication by a scales standard deviation by $|a|$.

6. Comparing Distributions

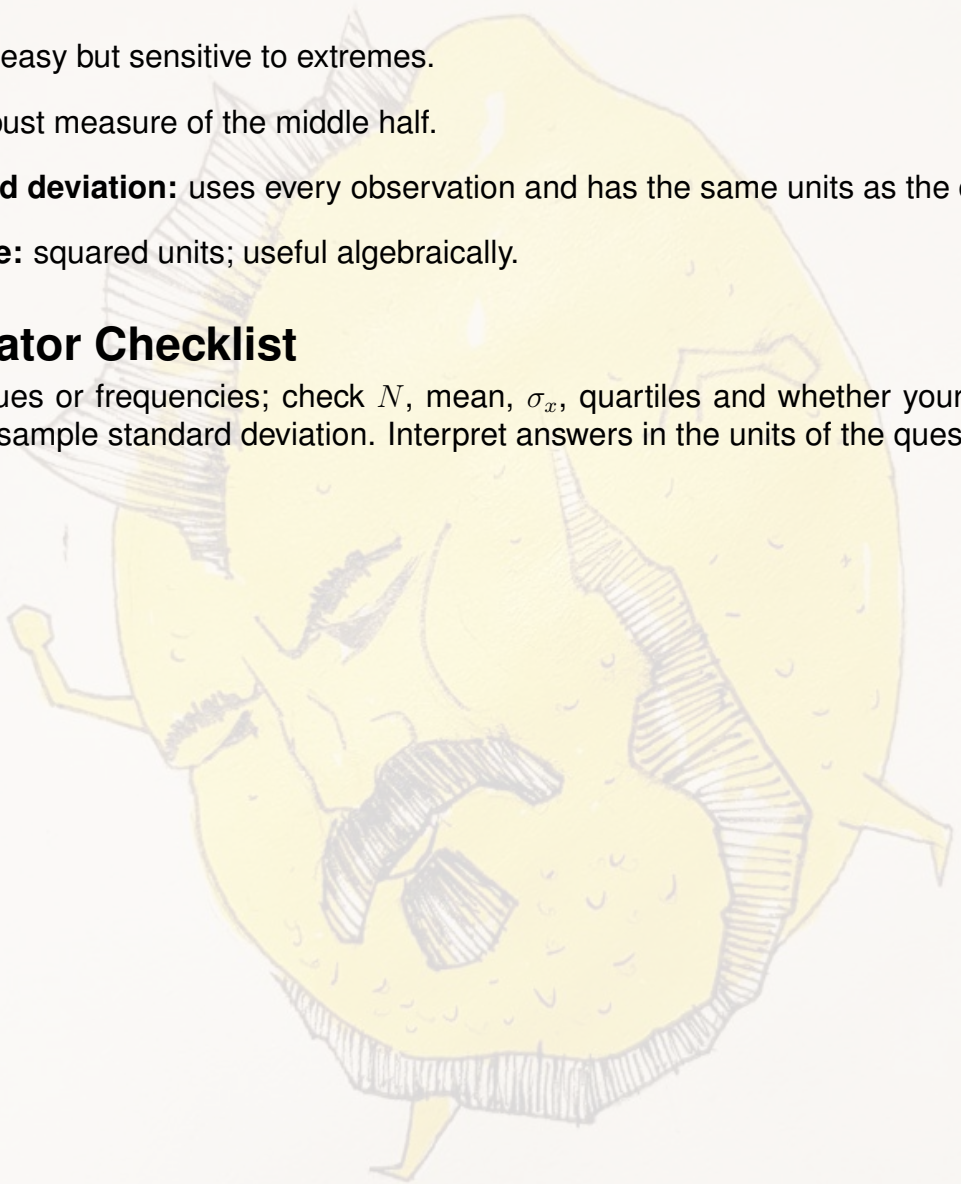
Use both centre and spread. A lower mean travel time is desirable, but a larger standard deviation means less consistency. For skewed data or outliers, median and IQR may be more informative than mean and standard deviation.

7. Choosing a Measure

- **Range:** easy but sensitive to extremes.
- **IQR:** robust measure of the middle half.
- **Standard deviation:** uses every observation and has the same units as the data.
- **Variance:** squared units; useful algebraically.

8. Calculator Checklist

Enter raw values or frequencies; check N, mean, σ_x, quartiles and whether your calculator uses population or sample standard deviation. Interpret answers in the units of the question.





Maths with Lemon

Mathematics made clear.