Statistics Unit 1 worksheet: Shifting and rescaling a data set (answer key)

ANSWERS AT A GLANCE

1. (a) 23 (b) 3.5 (c) 40 (d) 36 (e) 20 (f) 7 (g) 6 (h) 3
2. Mean 1165 m, median 1160 m, standard deviation 44 m, IQR 80 m.
3. (a) Mean 16.5 lb, standard deviation 3.08 lb. (b) Farm A is heavier on average (18 against 16.5 lb), and Farm A's weights also vary more (3.3 against 3.08 lb).
4. (a) $y = 15 + 4x$ (b) mean 63 s, median 51 s, standard deviation 20 s, variance 400 square seconds. (c) Every job gets the same 15 s, so it moves every time equally and no distance between times changes.
5. (a) Before: mean 40 dollars, range 20 dollars. After: mean 42.5 dollars, range 15 dollars. (b) The fee was not added to every value, so it is not a shift: only two of four bills rose, and the lowest bill rose while the highest did not.
6. Mean 75, median 65, standard deviation 15, variance 225, IQR 15.
7. (a) No: the standard deviation says $b = 2$, but the IQR says $b = 1.5$. (b) $b = 2$, $a = -10$.
8. Mean 1008 g; population standard deviation about 3.03 g.

WORKED SOLUTIONS

1. Each cell gives one summary of a data set x and the rule applied to every value. Find the summary asked for after the rule.

   (a) $\bar{x} = 14$, $y = x + 9$; find $\bar{y}$ (b) $\sigma_x = 3.5$, $y = x - 20$; find σ_y (c) median 8, $y = 5x$; find the new median (d) $\sigma_x^2 = 4$, $y = 3x$; find σ_y^2

   (e) IQR 10, $y = 2x + 7$; find the new IQR (f) $Q_1 = 6$, $y = 0.5x + 4$; find the new Q_1 (g) $\sigma_x = 1.5$, $y = 4x - 3$; find σ_y (h) $\sigma_x^2 = 2.25$, $y = 2x + 1$; find σ_y

   ANSWER
   (a) 23 (b) 3.5 (c) 40 (d) 36 (e) 20 (f) 7 (g) 6 (h) 3

   SOLUTION
   (a) A center moves by the constant: 14 + 9 = 23.
   (b) Taking away 20 moves every value the same distance, so the standard deviation stays 3.5.
   (c) A position is multiplied by 5: 5 × 8 = 40.
   (d) The variance is multiplied by $3^2 = 9$: 9 × 4 = 36.
   (e) A spread feels only the multiplier: 2 × 10 = 20. The 7 cancels.
   (f) A quartile goes through the whole rule: 0.5 × 6 + 4 = 7.
   (g) Only the multiplier acts on a spread: 4 × 1.5 = 6.
   (h) First $\sigma_x = \sqrt{2.25} = 1.5$, then $\sigma_y = 2 \times 1.5 = 3$. (Or: $\sigma_y^2 = 4 \times 2.25 = 9$, so $\sigma_y = 3$.)

2. A hiking trail has signposts along the way. Measured from the trailhead, the signposts stand at heights with mean 265 m, median 260 m, standard deviation 44 m and IQR 80 m. The trailhead itself is 900 m above sea level. Give the same four summaries for the signposts' heights above sea level.

   ANSWER
   Mean 1165 m, median 1160 m, standard deviation 44 m, IQR 80 m.

   SOLUTION
   Each height above sea level is the height above the trailhead plus 900 m: $y = x + 900$.
   Centers move by 900: mean 265 + 900 = 1165 m, median 260 + 900 = 1160 m.
   Spreads do not move, because every signpost rises by the same amount and the distances between them stay the same: standard deviation 44 m, IQR 80 m.

Free sample from math.wiki for classroom use. Do not post. math.wiki

3. Two farms weighed their watermelons. Farm A reports in pounds: mean 18 lb, standard deviation 3.3 lb. Farm B reports in kilograms: mean 7.5 kg, standard deviation 1.4 kg. Use 1 kg = 2.2 lb.

(a) Give Farm B's mean and standard deviation in pounds.

(b) Which farm's watermelons are heavier on average, and which farm's weights vary more?

ANSWER

(a) Mean 16.5 lb, standard deviation 3.08 lb. (b) Farm A is heavier on average (18 against 16.5 lb), and Farm A's weights also vary more (3.3 against 3.08 lb).

SOLUTION

(a) Each weight in pounds is the weight in kilograms times 2.2, a rescale with $k = 2.2$. Mean $2.2 \times 7.5 = 16.5$ lb; standard deviation $2.2 \times 1.4 = 3.08$ lb.

(b) Compare in the same unit. Farm A's mean 18 lb is above Farm B's 16.5 lb, and Farm A's standard deviation 3.3 lb is above Farm B's 3.08 lb. Compared in the raw numbers, 1.4 looks far smaller than 3.3; the units hid that the two spreads are close.

4. A printer needs 15 seconds to warm up for each print job and then 4 seconds for each page. One day its jobs had a mean of 12 pages, a median of 9 pages and a standard deviation of 5 pages.

(a) Write the time y, in seconds, for a job of x pages.

(b) Find the mean, median, standard deviation and variance of the job times.

(c) Explain why the 15 seconds of warm-up plays no part in the standard deviation.

ANSWER

(a) $y = 15 + 4x$ (b) mean 63 s, median 51 s, standard deviation 20 s, variance 400 square seconds. (c) Every job gets the same 15 s, so it moves every time equally and no distance between times changes.

SOLUTION

(a) Each job takes 15 s plus 4 s per page: $y = 15 + 4x$, a rule $a + bx$ with $a = 15$, $b = 4$.

(b) Mean $15 + 4 \times 12 = 63$ s; median $15 + 4 \times 9 = 51$ s. Standard deviation $4 \times 5 = 20$ s; variance $20^2 = 400$ square seconds.

(c) The deviation of a job time from the mean time is $(15 + 4x) - (15 + 4\bar{x}) = 4(x - \bar{x})$: the 15 cancels. Every job waits the same 15 s, so the times are all pushed along by the same amount and their spread is unchanged.

5. A tutoring center sent four families bills of 30, 35, 45 and 50 dollars. It then added a 5 dollar fee to every bill under 40 dollars and nothing to the others.

(a) Find the mean and the range of the bills before and after the fee.

(b) Explain why the new mean is not the old mean plus 5, and why the range changed.

ANSWER

(a) Before: mean 40 dollars, range 20 dollars. After: mean 42.5 dollars, range 15 dollars. (b) The fee was not added to every value, so it is not a shift: only two of four bills rose, and the lowest bill rose while the highest did not.

SOLUTION

(a) Before: $\dfrac{30 + 35 + 45 + 50}{4} = \dfrac{160}{4} = 40$ dollars; range $50 - 30 = 20$ dollars.

After: the bills are 35, 40, 45, 50. Mean $\dfrac{170}{4} = 42.5$ dollars; range $50 - 35 = 15$ dollars.

Free sample from math.wiki for classroom use. Do not post. math.wiki

(b) The rules of this lesson need the same change on every value. Here only two of the four bills gained 5, so the total grew by 10, not 20, and the mean by $\frac{10}{4} = 2.5$. The fee raised the smallest bill but not the largest, so the bills moved closer together and the range shrank by 5.

6. Samir gave a quiz marked out of 20 points. To put the scores out of 100, every score is multiplied by 5, and then everyone gets 5 bonus points. Before the change the scores had mean 14, median 12, standard deviation 3 and IQR 3 points. Find the new mean, median, standard deviation, variance and IQR.

ANSWER
Mean 75, median 65, standard deviation 15, variance 225, IQR 15.

SOLUTION
The rule is $y = 5x + 5$.

Centers go through the whole rule: mean 5 × 14 + 5 = 70 + 5 = 75, median 5 × 12 + 5 = 60 + 5 = 65.

Spreads feel only the multiplier: standard deviation 5 × 3 = 15, IQR 5 × 3 = 15. The variance is the square of the new standard deviation, $15^2 = 225$ (the old variance times 5^2).

7. Malik says that every score on a test was changed by one rule $y = a + bx$ with $b > 0$. Before the change the scores had mean 40, standard deviation 5 and IQR 8. After it they had mean 70, standard deviation 10 and IQR 12.

(a) Could Malik be right? Explain.

(b) Suppose the IQR after the change had been 16 instead. Find a and b.

ANSWER
(a) No: the standard deviation says $b = 2$, but the IQR says $b = 1.5$. (b) $b = 2$, $a = -10$.

SOLUTION
(a) Under $y = a + bx$ every spread is multiplied by the same b. The standard deviation went from 5 to 10, so $b = 2$. The IQR went from 8 to 12, so $b = 1.5$. One rule cannot have two multipliers, so the scores were not all changed by one rule of this kind.

(b) Now $16 = 8b$ also gives $b = 2$, which agrees with the standard deviation. The mean gives a: $70 = a + 2 \times 40$, so $a = -10$. The rule is $y = 2x - 10$.

8. Five bags of flour are weighed, in grams: 1003, 1007, 1008, 1010, 1012. Subtract 1000 from every weight first, and use the smaller numbers to find the mean and the population standard deviation of the five weights. Round the standard deviation to two decimal places.

ANSWER
Mean 1008 g; population standard deviation about 3.03 g.

SOLUTION
After subtracting 1000 the values are 3, 7, 8, 10, 12. Their mean is $\frac{40}{5} = 8$. The deviations are $-5, -1, 0, 2, 4$, the squares 25, 1, 0, 4, 16, with sum 46. The variance is $\frac{46}{5} = 9.2$ and the standard deviation $\sqrt{9.2} \approx 3.03$.

Subtracting 1000 moved the mean down by 1000 and left the standard deviation alone, so for the real weights the mean is 8 + 1000 = 1008 g and the standard deviation is about 3.03 g.

Free sample from math.wiki for classroom use. Do not post. math.wiki