Statistics Unit 1 quiz 3: Resistant and non resistant measures, Box plots that show outliers, Comparing two distributions (answer key)

7 questions, 21 points

ANSWERS AT A GLANCE

1. Mean 22 kg, median 14 kg. She should use the median, 14 kg.
2. (a) Mean 20 dollars, median 16 dollars. (b) The median, 16 dollars. (c) The mean: 20 × 8 × 30 = 4,800 dollars.
3. (a) 25 and 65. (b) Yes on both sides: the minimum 12 is below 25 and the maximum 95 is above 65. (c) No. It ends at the largest time that is at most 65, which the summary does not give; it is somewhere from 50 to 65.
4. (a) Brand A: median 31 raisins, range 6 raisins. Brand B: median 30 raisins, range 13 raisins. (b) "In these samples, the median Brand A box, 31 raisins, holds 1 raisin more than the median Brand B box, 30 raisins." "Brand B's counts are far more spread out, with a range of 13 raisins against 6." (c) Brand A: 0 boxes; Brand B: 3 boxes. The typical boxes are close, but Brand B is less consistent, so its buyer is more likely to get a box with fewer than 28.
5. (a) Mean 34 people, MAD 10 people. (b) Mean 40 people, MAD 2.4 people. (c) About 4.2 times larger.
6. Five-number summary: 38, 54.5, 58.5, 61.5, 66. Fences: 44 and 72. Outlier: 38. The plot: a box from 54.5 to 61.5 with a line at 58.5; the whiskers end at 52 and 66; one dot at 38. No fence is drawn.
7. (a) The means, with the ranges, for both farms. (b) The medians and interquartile ranges, for both shops. (c) No single center fits Shop D; describe its two clusters separately and compare each with Shop C.

WORKED SOLUTIONS

1. Ava's class collects bottles and cans for recycling. In eight weeks it collected 14, 11, 16, 12, 15, 13, 17, 10 kg. In the ninth week the school cleared out its storeroom, and the class collected 90 kg.

Find the mean and the median of all nine weeks. Ava's teacher wants to tell the principal how much a usual week brings in. Which of the two should she use? Explain. (3 points)

> **ANSWER**
> Mean 22 kg, median 14 kg. She should use the median, 14 kg.
>
> **SOLUTION**
>
> Sorted: 10, 11, 12, 13, 14, 15, 16, 17, 90. The total is 108 + 90 = 198, so the mean is $\frac{198}{9} = 22$ kg. Nine values, so the median is the fifth, 14 kg.
>
> The mean of 22 kg is more than the class collected in any of the eight ordinary weeks, because the one storeroom week added 90 kg to the total that is shared among all nine weeks. The median only uses the middle place of the sorted list, so the 90 barely moves it, and 14 kg sits among the usual weeks.
>
> **SCORING**
> 1 point: mean 22 kg. 1 point: median 14 kg. 1 point: chooses the median with a reason that names the one large week pulling the mean above every usual week.

2. The eight staff at a summer day camp earn these hourly pay rates, in dollars: the seven counselors earn 15, 16, 15, 17, 16, 18, 15, and the camp director earns 48. (3 points)

(a) Find the mean and the median hourly pay of the eight staff.

(b) The camp's website will give one number for what a staff member usually earns per hour. Which of the two should it use?

(c) Each of the eight staff works 30 hours this week. Which of the two centers tells the camp the total pay for the week, and what is that total?

> **ANSWER**

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

(a) Mean 20 dollars, median 16 dollars. (b) The median, 16 dollars. (c) The mean: 20 × 8 × 30 = 4,800 dollars.

SOLUTION

(a) Sorted: 15, 15, 15, 16, 16, 17, 18, 48. The total is 160, so the mean is $\frac{160}{8} = 20$ dollars. The median is halfway between the fourth and fifth values, $\frac{16 + 16}{2} = 16$ dollars.

(b) The median. Seven of the eight staff earn 18 dollars or less, so a mean of 20 dollars is more than any counselor earns; the director's pay pulled it up. The median, 16 dollars, is a usual rate.

(c) The mean. The total hourly pay is the mean times the count, 20 × 8 = 160 dollars per hour for the whole staff, so the week costs 160 × 30 = 4,800 dollars. The median times the count, 16 × 8 × 30 = 3,840 dollars, would leave out most of the director's pay.

SCORING

1 point: mean 20 dollars and median 16 dollars. 1 point: median for the website, with a reason. 1 point: names the mean for the total and finds 4,800 dollars.

3. The members of a puzzle club timed how many minutes they took to finish one crossword. The only numbers kept were the five-number summary: minimum 12, $Q_1 = 40$, median 45, $Q_3 = 50$, maximum 95. (3 points)

(a) Find the fences.

(b) Is the modified box plot sure to have a dot below the box? Above it? Explain.

(c) Can you tell where its upper whisker ends? Explain.

ANSWER

(a) 25 and 65. (b) Yes on both sides: the minimum 12 is below 25 and the maximum 95 is above 65. (c) No. It ends at the largest time that is at most 65, which the summary does not give; it is somewhere from 50 to 65.

SOLUTION

(a) IQR = 50 − 40 = 10 and 1.5 × 10 = 15, so the fences are 40 − 15 = 25 and 50 + 15 = 65.

(b) The minimum, 12, is a real time below 25, so it is a low outlier and gets a dot. The maximum, 95, is a real time above 65, so it is a high outlier and gets a dot.

(c) No. The upper whisker ends at the largest time that is not an outlier, the largest one at most 65. The summary gives only $Q_3 = 50$ and the maximum 95; the times between are unknown. The whisker end is somewhere from 50 to 65, but which time it is cannot be told.

SCORING

1 point for the fences 25 and 65; 1 point for a dot on both sides, with the minimum and maximum compared to the fences; 1 point for saying the whisker end cannot be found, because the times between Q_3 and the maximum are unknown.

4. Yuna bought random samples of 11 small boxes of raisins of each of two brands and counted the raisins in each box.

Brand A: 28, 29, 30, 30, 31, 31, 31, 32, 32, 33, 34

Brand B: 24, 26, 27, 28, 29, 30, 30, 31, 33, 35, 37 (3 points)

(a) Find the median and the range of each brand.

(b) Write one comparison sentence about center and one about spread.

(c) How many boxes of each brand hold fewer than 28 raisins? What does the difference in spread mean for a buyer who wants at least 28 raisins?

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

ANSWER

(a) Brand A: median 31 raisins, range 6 raisins. Brand B: median 30 raisins, range 13 raisins. (b) "In these samples, the median Brand A box, 31 raisins, holds 1 raisin more than the median Brand B box, 30 raisins." "Brand B's counts are far more spread out, with a range of 13 raisins against 6." (c) Brand A: 0 boxes; Brand B: 3 boxes. The typical boxes are close, but Brand B is less consistent, so its buyer is more likely to get a box with fewer than 28.

SOLUTION

(a) With 11 values the median is the 6th. Brand A: median 31, range 34 − 28 = 6 raisins. Brand B: median 30, range 37 − 24 = 13 raisins.

(b) "In these samples, the median Brand A box, 31 raisins, holds 1 raisin more than the median Brand B box, 30 raisins." "Brand B's counts are far more spread out, with a range of 13 raisins against 6 raisins."

(c) Brand A's smallest count is 28, so no box has fewer. Brand B has 24, 26 and 27: 3 of its 11 boxes. The centers differ by only 1 raisin, but Brand B's wider spread puts more boxes far from its center, on both sides. A buyer who wants at least 28 was never short with Brand A in this sample and was short about 3 times in 11 with Brand B.

SCORING

1 point for both medians and both ranges; 1 point for two comparison sentences that name the measure, both values with units and the direction and size of the difference; 1 point for 0 and 3 boxes with the link to Brand B's larger spread.

5. A community yoga class meets in the park every Saturday. The numbers of people at the last six classes were 40, 44, 36, 4, 42, 38. On the day only 4 came, the class was moved indoors at the last minute. (3 points)

(a) Find the mean and the MAD of all six classes.

(b) Find the mean and the MAD of the five classes left when the day with 4 people is left out.

(c) How many times larger is the MAD with that day than without it?

ANSWER

(a) Mean 34 people, MAD 10 people. (b) Mean 40 people, MAD 2.4 people. (c) About 4.2 times larger.

SOLUTION

(a) The total is 204, so the mean is $\frac{204}{6} = 34$ people. The distances from 34 are 6, 10, 2, 30, 8, 4, total 60, so the MAD is $\frac{60}{6} = 10$ people.

(b) The total is 200, so the mean is $\frac{200}{5} = 40$ people. The distances from 40 are 0, 4, 4, 2, 2, total 12, so the MAD is $\frac{12}{5} = 2.4$ people.

(c) $\frac{10}{2.4} \approx 4.2$, so the MAD is about 4.2 times larger with the indoor day. That one day's distance of 30 alone adds $\frac{30}{6} = 5$ to the MAD, and it also pulled the mean down to 34, away from the five ordinary classes.

SCORING

1 point: mean 34 and MAD 10. 1 point: mean 40 and MAD 2.4. 1 point: about 4.2 times.

6. A farm weighed 12 eggs, in grams: 38, 52, 54, 55, 57, 58, 59, 60, 61, 62, 63, 66. Draw the modified box plot of the weights on a number line from 35 to 70, and show the work that places each part of it. (3 points)

ANSWER

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

Five-number summary: 38, 54.5, 58.5, 61.5, 66. Fences: 44 and 72. Outlier: 38. The plot: a box from 54.5 to 61.5 with a line at 58.5; the whiskers end at 52 and 66; one dot at 38. No fence is drawn.

SOLUTION

With 12 values the median is $\frac{58 + 59}{2} = 58.5$. The lower half 38, 52, 54, 55, 57, 58 gives $Q_1 = \frac{54 + 55}{2} = 54.5$, and the upper half 59, 60, 61, 62, 63, 66 gives $Q_3 = \frac{61 + 62}{2} = 61.5$.

IQR = 61.5 − 54.5 = 7 and 1.5 × 7 = 10.5, so the fences are 54.5 − 10.5 = 44 and 61.5 + 10.5 = 72.

Only 38 is below 44, so it is an outlier with its own dot, and the lower whisker stops at the smallest egg that is not an outlier, 52. No egg is above 72, so the upper whisker reaches the maximum, 66.

SCORING

1 point for the five-number summary; 1 point for the fences 44 and 72 with 38 named as the only outlier; 1 point for the drawing: box from 54.5 to 61.5 with the median line at 58.5, lower whisker stopping at 52 (not at 44 or 38), upper whisker to 66, and a dot at 38.

7. For each pair of samples, say which measures of center and spread Hiro should use to compare the two groups, and why. (3 points)

(a) The lengths of random samples of 25 cucumbers from each of two farms. Both distributions are roughly symmetric with one peak and no outliers.

(b) The minutes that random samples of 30 shoppers spent in two shops. Shop 1's times are roughly symmetric. Shop 2's times are skewed right, with two shoppers far above all the rest.

(c) The prices of random samples of 20 bikes at two shops. Shop C's prices are roughly symmetric around 300 dollars. Shop D's prices form two clusters, children's bikes from 80 to 150 dollars and adult bikes from 350 to 600 dollars, with no prices in between.

ANSWER

(a) The means, with the ranges, for both farms. (b) The medians and interquartile ranges, for both shops. (c) No single center fits Shop D; describe its two clusters separately and compare each with Shop C.

SOLUTION

(a) Both shapes are roughly symmetric with no outliers, so the mean, which uses every value, describes each center well; compare the means, with the range for spread.

(b) Shop 2's skew and its two far-out shoppers would pull its mean toward the tail, while the median moves only when the middle of the list moves. Shop 2 calls for the median, and one measure for both groups means medians and interquartile ranges for both shops, even though Shop 1 alone could use the mean. A mean set against a median compares two different things.

(c) With two clusters, Shop D's mean or median may land in the gap between 150 and 350 dollars, where no bike is priced. Describe each cluster: the children's bikes and the adult bikes, each with its own center and spread, and say where Shop C's prices, near 300 dollars, sit against each cluster.

SCORING

1 point for each part: (a) means and ranges, because both shapes are symmetric with no outliers; (b) medians and interquartile ranges for both shops, because Shop 2 is skewed with far-out values and the same measure is used for both; (c) no single center for Shop D, because it would fall in the gap, so describe each cluster. A part with the right measures but no reason from the shape earns no point.

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