Statistics Unit 1 worksheet: Describing a distribution (answer key)

ANSWERS AT A GLANCE

1. (a) Sentence 1: context. Sentence 2: shape. Sentence 3: center and spread. Sentence 4: the answer to the question. (b) Unusual features. (c) The lower fence is 6.25 cm and $6 < 6.25$, so the stub is an outlier. For example: "The 6 cm stub is an outlier; no pencil is long enough to be one at the top."
2. (a) Confirmed. (b) Confirmed. (c) Refuted: the lower fence is 6.5 cm. (d) Confirmed: the fences are 6.5 and 26.5 cm, and the minimum and maximum are inside them. (e) Cannot be checked. (f) Cannot be checked.
3. (a) Mean 20, MAD 0.8, median 20, $Q_1 = 19$, $Q_3 = 21$; fences 16 and 24, no outliers. (b) Roughly symmetric, so the mean and the MAD: "Mason counted the sheets in 10 notepads from one pack. The counts are roughly symmetric, with one peak at 20, and run from 18 to 22 with no gaps and no outliers. The mean is 20 sheets, and the notepads are on average 0.8 sheets away from it. A notepad from this pack holds about 20 sheets." (c) The median and the IQR: median 20, IQR $21 - 19 = 2$ sheets; the upper fence is 24, so the 34 is an outlier.
4. (a) The fences are 24.75 and 34.75 pieces: the 19 is an outlier and the 33 is not. (b) The median with the IQR, because of the outlier. (c) "Grace counted the fruit snacks in 17 fun-size bags. The counts are skewed left by one low bag; the other 16 are roughly symmetric around 30, from 27 to 33 pieces. The median is 30 pieces, and the middle half of the bags hold 28.5 to 31 pieces, an IQR of 2.5 pieces. One bag, with only 19 pieces, is an outlier, with a gap from 20 to 26 pieces where there are no bags. A fun-size bag typically holds about 30 fruit snacks." For the shape, either "skewed left" or "roughly symmetric apart from one low outlier" is right, as long as the 19 is named as the outlier that causes it.
5. (a) Both are right: the total is 49, so the mean is 2.45, and the distances total 33.7, so the MAD is 1.685, about 1.7 books. (b) The counts are skewed right, not symmetric, and the 11 is an outlier (the upper fence is 6). The pair is also wrong: skewed data with an outlier calls for the median and the IQR. (c) "Leah asked the 20 students in her class how many library books each checked out this month. The counts are skewed right. The median is 2 books, and the middle half of the students checked out 1 to 3 books, an IQR of 2 books. One student checked out 11 books, an outlier, with a gap from 7 to 10. A typical student in the class checked out about 2 books this month."

WORKED SOLUTIONS

1. Jin measured every pencil in his class's pencil cup and wrote this description, one sentence at a time.

- Sentence 1: "I measured the 18 pencils in our class's pencil cup with a ruler, to the nearest centimeter."
- Sentence 2: "The lengths are skewed left: most pencils are 14 to 19 cm long, and a few worn-down ones trail off to the shortest, a stub of 6 cm."
- Sentence 3: "The median length is 16 cm, and the middle half of the pencils are 13 to 17.5 cm long, an IQR of 4.5 cm."
- Sentence 4: "A typical pencil in the cup is about 16 cm long."

(a) For each sentence, name the piece of a description it gives (context, shape, center and spread, unusual features, or the answer to the question).

(b) Which piece is missing?

(c) Use the numbers in Sentence 3 to decide whether the 6 cm stub is an outlier. Then write the missing sentence.

> **ANSWER**
> (a) Sentence 1: context. Sentence 2: shape. Sentence 3: center and spread. Sentence 4: the answer to the question. (b) Unusual features. (c) The lower fence is 6.25 cm and $6 < 6.25$, so the stub is an outlier. For example: "The 6 cm stub is an outlier; no pencil is long enough to be one at the top."
>
> **SOLUTION**
> (a) Sentence 1 names the individuals (the 18 pencils), the variable and unit (length in centimeters) and how it was measured, so it is the context. Sentence 2 gives the shape. Sentence 3 gives a matching pair, the median with the IQR, which suits skewed data. Sentence 4 answers "How long are the pencils?" in the setting's words.

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

(b) No sentence says whether there are outliers, gaps or clusters: the unusual features are missing.

(c) The IQR is 17.5 − 13 = 4.5 cm and 1.5 × 4.5 = 6.75. The fences are 13 − 6.75 = 6.25 cm and 17.5 + 6.75 = 24.25 cm. The stub, 6 cm, is below 6.25, so it is an outlier. The longest pencil, 19 cm, is below 24.25, so there is no outlier at the top.

Missing sentence, for example: "The 6 cm stub is an outlier, below the lower fence of 6.25 cm; it may be a pencil that was used almost to the end."

2. Olivia's class measured the 36 carrots in a grocery crate, to the nearest centimeter, and kept only the five-number summary: minimum 8, $Q_1 = 14$, median 17, $Q_3 = 19$ and maximum 21 cm.

For each claim, write confirmed, refuted, or cannot be checked from the summary, and give the reason.

(a) "The middle half of the carrots are 14 to 19 cm long, an IQR of 5 cm."

(b) "The lengths are skewed left."

(c) "The shortest carrot, 8 cm, is an outlier."

(d) "There are no outliers at all."

(e) "No carrot is between 9 and 13 cm long."

(f) "The mean length is 16 cm."

ANSWER

(a) Confirmed. (b) Confirmed. (c) Refuted: the lower fence is 6.5 cm. (d) Confirmed: the fences are 6.5 and 26.5 cm, and the minimum and maximum are inside them. (e) Cannot be checked. (f) Cannot be checked.

SOLUTION

(a) The middle half runs from $Q_1 = 14$ to $Q_3 = 19$ cm, and 19 − 14 = 5 cm. Confirmed.

(b) The left whisker is 14 − 8 = 6 cm and the left half of the box 17 − 14 = 3 cm; the right half is 19 − 17 = 2 cm and the right whisker 21 − 19 = 2 cm. Each piece holds about a quarter of the carrots, and the left pieces are stretched longer, so the lengths trail off to the left. Confirmed.

(c) 1.5 × 5 = 7.5, so the lower fence is 14 − 7.5 = 6.5 cm. The 8 cm carrot is above it, so it is not an outlier. Refuted.

(d) The upper fence is 19 + 7.5 = 26.5 cm. Every carrot is between the minimum, 8, and the maximum, 21, and both are inside the fences, so no carrot can be an outlier. Confirmed.

(e) The whisker from 8 to 14 holds about a quarter of the carrots, but the summary does not show where they sit on it. They might all be 12 to 14 cm, or spread along it. Cannot be checked: a summary hides gaps.

(f) The five numbers do not give the total of the lengths, so the mean cannot be found. Cannot be checked.

3. Mason bought a big pack of small notepads, each sold as about 20 sheets. He counted the sheets in 10 of them: 20, 18, 21, 20, 19, 22, 20, 19, 21, 20.

(a) Find the mean, the MAD, the median and the quartiles, and test for outliers.

(b) Name the shape, choose a center and a spread, and write a description that answers "How many sheets are in a notepad from this pack?"

(c) Mason then finds an 11th notepad in the pack, made of much thinner paper, that holds 34 sheets. Which center and spread should the description use now? Find them, and test the 34 for being an outlier.

ANSWER

(a) Mean 20, MAD 0.8, median 20, $Q_1 = 19$, $Q_3 = 21$; fences 16 and 24, no outliers. (b) Roughly symmetric, so the mean and the MAD: "Mason counted the sheets in 10 notepads from one pack. The counts are roughly symmetric, with one peak at 20, and run from 18 to 22 with no gaps and no outliers. The mean is 20 sheets, and the notepads are on average 0.8 sheets away from it. A notepad from this pack

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

holds about 20 sheets." (c) The median and the IQR: median 20, IQR 21 − 19 = 2 sheets; the upper fence is 24, so the 34 is an outlier.

SOLUTION

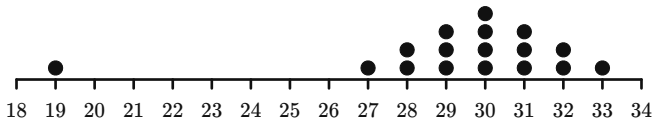
(a) Sorted: 18, 19, 19, 20, 20, 20, 20, 21, 21, 22. The total is 200, so the mean is $\frac{200}{10} = 20$. The distances from 20 are 2, 1, 1, 0, 0, 0, 0, 1, 1, 2, which total 8, so the MAD is $\frac{8}{10} = 0.8$. The median is $\frac{20 + 20}{2} = 20$. The lower half 18, 19, 19, 20, 20 gives $Q_1 = 19$ and the upper half 20, 20, 21, 21, 22 gives $Q_3 = 21$. The IQR is 2 and 1.5 × 2 = 3, so the fences are 16 and 24. Every count is between them: no outliers.

(b) One peak at 20, with 3 notepads below it and 3 above it, at the same distances. The mean equals the median. Roughly symmetric with no outliers, so the mean with the MAD.

Description: "Mason counted the sheets in 10 notepads from one pack. The counts are roughly symmetric, with one peak at 20, and run from 18 to 22 with no gaps and no outliers. The mean is 20 sheets, and the notepads are on average 0.8 sheets away from it (the MAD). A notepad from this pack holds about 20 sheets."

(c) Sorted: 18, 19, 19, 20, 20, 20, 20, 21, 21, 22, 34. With 11 values the median is the 6th, 20. The lower half 18, 19, 19, 20, 20 gives $Q_1 = 19$ and the upper half 20, 21, 21, 22, 34 gives $Q_3 = 21$, so the IQR is 2 and the upper fence is 21 + 3 = 24. The 34 is beyond it: an outlier. The mean is now $\frac{234}{11}$, about 21.3, above 9 of the 11 notepads, so it no longer shows a typical notepad. Use the resistant pair: the median, 20 sheets, with the IQR, 2 sheets, and name the 34 as an outlier.

4. Grace opened 17 fun-size bags of fruit snacks and counted the pieces in each bag. Her summaries: mean about 29.4, median 30, $Q_1 = 28.5$ and $Q_3 = 31$ pieces.



Fruit snacks in one bag

(a) Test the bag with 19 pieces and the bag with 33 pieces for being outliers.

(b) Which center and spread should Grace's description use? Explain.

(c) Write a description that answers "How many fruit snacks are in a fun-size bag?" Include any gap.

ANSWER

(a) The fences are 24.75 and 34.75 pieces: the 19 is an outlier and the 33 is not. (b) The median with the IQR, because of the outlier. (c) "Grace counted the fruit snacks in 17 fun-size bags. The counts are skewed left by one low bag; the other 16 are roughly symmetric around 30, from 27 to 33 pieces. The median is 30 pieces, and the middle half of the bags hold 28.5 to 31 pieces, an IQR of 2.5 pieces. One bag, with only 19 pieces, is an outlier, with a gap from 20 to 26 pieces where there are no bags. A fun-size bag typically holds about 30 fruit snacks." For the shape, either "skewed left" or "roughly symmetric apart from one low outlier" is right, as long as the 19 is named as the outlier that causes it.

SOLUTION

(a) The IQR is 31 − 28.5 = 2.5 pieces and 1.5 × 2.5 = 3.75. The fences are 28.5 − 3.75 = 24.75 and 31 + 3.75 = 34.75 pieces. The 19 is below 24.75, so it is an outlier. The 33 is below 34.75, so it is not: it only looks a little far.

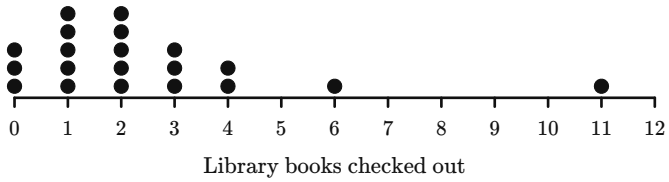
(b) The other 16 bags are roughly symmetric around 30, but the outlier at 19 pulls the mean down to about 29.4, below the median. The median and the IQR are resistant, so they stay with the bulk of the bags. Keep them together: the MAD would measure distances from the mean.

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

(c) The plot has no dots at 20, 21, ..., 26: a gap between the outlier and the rest. Shape: sixteen bags sit between 27 and 33 and one stands far off to the left, and the mean is below the median, so "skewed left" fits, as with the one slow sled run in the lesson. Saying the other 16 are roughly symmetric is also true; the description says both.

Description: "Grace counted the fruit snacks in 17 fun-size bags. The counts are skewed left by one low bag; the other 16 are roughly symmetric around 30, from 27 to 33 pieces. The median is 30 pieces, and the middle half of the bags hold 28.5 to 31 pieces, an IQR of 2.5 pieces. One bag, with only 19 pieces, is an outlier, with a gap from 20 to 26 pieces where there are no bags; it may have been filled wrongly at the factory. A fun-size bag typically holds about 30 fruit snacks."

5. Leah asked the 20 students in her class how many library books each one checked out this month, and drew the dot plot. Hiro wrote this description of her data:



"The 20 students checked out a mean of 2.45 books, with a MAD of about 1.7 books. The counts are roughly symmetric, and there are no outliers."

(a) Use the dot plot to check Hiro's mean and MAD.

(b) Which of Hiro's claims are wrong? Use the shape of the plot and the fences.

(c) Write a better description that answers "How many library books does a student in Leah's class check out in a month?"

ANSWER

(a) Both are right: the total is 49, so the mean is 2.45, and the distances total 33.7, so the MAD is 1.685, about 1.7 books. (b) The counts are skewed right, not symmetric, and the 11 is an outlier (the upper fence is 6). The pair is also wrong: skewed data with an outlier calls for the median and the IQR. (c) "Leah asked the 20 students in her class how many library books each checked out this month. The counts are skewed right. The median is 2 books, and the middle half of the students checked out 1 to 3 books, an IQR of 2 books. One student checked out 11 books, an outlier, with a gap from 7 to 10. A typical student in the class checked out about 2 books this month."

SOLUTION

(a) Total: 0 × 3 + 1 × 5 + 2 × 5 + 3 × 3 + 4 × 2 + 6 + 11 = 49, so the mean is $\frac{49}{20} = 2.45$ books. Distances from 2.45: the three 0s give 3 × 2.45 = 7.35, the five 1s 5 × 1.45 = 7.25, the five 2s 5 × 0.45 = 2.25, the three 3s 3 × 0.55 = 1.65, the two 4s 2 × 1.55 = 3.1, then 3.55 and 8.55. The total is 33.7 and $\frac{33.7}{20} = 1.685$, about 1.7 books. Both numbers are right.

(b) The dots crowd at 0 to 2, against the floor of 0 books, and thin out to the right as far as 11: skewed right, not symmetric. The mean, 2.45, above the median, agrees. Sorted, the 10th and 11th counts are both 2, so the median is 2. The lower ten give $Q_1 = \frac{1+1}{2} = 1$ and the upper ten $Q_3 = \frac{3+3}{2} = 3$. The IQR is 2, 1.5 × 2 = 3, and the upper fence is 3 + 3 = 6. The 11 is beyond it, so "no outliers" is false; the 6 sits on the fence and is not an outlier. With the shape and the outlier, the mean and MAD are the wrong pair.

(c) "Leah asked the 20 students in her class how many library books each checked out this month. The counts are skewed right. The median is 2 books, and the middle half of the students checked out 1 to 3

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

> books, an IQR of 2 books. One student checked out 11 books, an outlier, with a gap from 7 to 10 books where no student is. A typical student in the class checked out about 2 books this month." (No student checked out 5 books either; that one empty count inside the tail may be mentioned, but the gap that matters is the one before the outlier.)

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