CENTER AND SPREAD HOMEWORK

TEXT: 2.5, 2.6, 2.7

LAST NAME	FIRST NAME	DATE

For each of the following samples, find (a) the mean, (b) the median, (c) the range, which is min – max, (d) the variance, and (e) the standard deviation. You do not need to show any work here, but **you should compute the variance by hand, using a formula or a table**. You can use R as a basic calculator, but you should not employ any R functions beyond `sqrt`. Once you've computed the statistics by hand, you are welcome to check your answers using R's functions such as `mean` and `sd`.

1. `8, -3, 9, 5, 1`

2. `-10, -13, -14, -18, -20`

3. `-7, -6, 9, 2, -4, 0`

4. `0, 9, -6, -8, 5, 4, 10`

5. `4, -5, 4, -4, -5, -2, 3, -3`

6. `17, 17, 17, 17`

Use R to find the following statistics for the dataset `LakeHuron`

7. Mean, with the units of measurement

8. Median

9. Mode(s)

10. Minimum

11. Maximum

12. Variance

13. Standard deviation

14. Median absolute deviation (MAD)

Given a sample with mean $\overline{x} = -42$ and standard deviation $s = 9$, find the following:

15. Compute the z-score for the data value -17.

16. Compute the z-score for the data value 3.14.

17. Find the data value with z-score -4.

18. Assuming that the distribution is bell-shaped and symmetric, which one of the above data values is most likely? Which one is least likely?

Given a population with mean $\mu = -10$ and standard deviation $\sigma = 6$, and assuming that the distribution is **approximately normal**, find the following:

19. What percentage of measurements would you expect to lie between -28 and -10?

20. What percentage of measurements would you expect to lie between -16 and 2?

21. What percentage of measurements would you expect to lie above 8?

22. What percentage of measurements would you expect to lie below -4?

Assume still that $\mu = -10$ and $\sigma = 6$ but **do not assume anything about the shape of the population distribution** as you answer the following questions.

23. What percentage of measurements would you expect to be between -20 and 0?

24. What percentage of measurements would you expect to be between -40 and 20?

Use R to find the following statistics for the dataset `state.area`

25. Mean, with the units of measurement

26. Median

27. Mode(s)

28. Range, which is min – max

29. Variance

30. Standard deviation

31. MAD

32. Mean absolute deviation from the mean

33. Mean absolute deviation from the median

34. Median absolute deviation from the mean

Note that you can use the built-in `mad` function to compute the median absolute deviation from the median, but you can also compute it manually for any dataset `x`:

```
median(abs(x - median(x)))
```

You can use a similar solution for the last 3 problems.

35. Use the Chebyshev's inequality to estimate the proportion of US states with areas more than 3 standard deviations away from the mean.

36. Consult the `state.area` dataset to see what is the actual proportion of the US states with areas more than 3 standard deviations away from the mean.

ANSWERS

1.

(a) $\overline{x} = 4$
(b) median $= 5$
(c) range $= 12$
(d) $s^2 = 25$
(e) $s = 5$

3.

(a) $\overline{x} = -1$
(b) median $= -2$
(c) range $= 16$
(d) $s^2 = 36$
(e) $s = 6$

5.

(a) $\overline{x} = -1$
(b) median $= -2.5$
(c) range $= 9$
(d) $s^2 = 16$
(e) $s = 4$

7. 579.0041 feet

9. Judging by the histogram, about 579.5 feet

11. 581.86 feet

13. 1.318299 feet

15. -2.777778

17. -78

19. Approximately 50%

21. Less than 1%

23. At least 64%

25. 72367.98 square miles

27. About 50000 square miles

29. 7793006379 mi^4

31. 24087 square miles

33. 42194.7 square miles

35. At most 11.11%