

Name: Nasser **Al Mufarji**

Use the following story to answer questions 1-5:

The Norwegian Soap Company, Hans R. Derty, decided to do a market analysis on the preferences of farmers. They randomly selected 150 farmers. The mean preference score for the soap used by the farmers was 65. From this it was determined that the mean preference score would be 62-68 for all Norwegian farmers. Given these results, they decided that grime didn't pay!

1. What is the sample? (1 pt.)

Randomly selected 150 farmers

2. What is the parameter? (1 pt.)

It is between 62 - 68

3. What is the statistic? (1 pt.)

The mean average score of 65

4. What is the population? (1 pt.)

All farmers at the Norwegian Soap Company

5. The numeral 150 in the above situation is used as what type of data and what type of measurement scale?

Data: It is discrete (1 pt.)

Measurement: Ratio (1 pt.)

6. Lois Carmen Denominator obtained $P_{81} = 63$. How do you interpret this? (1 pt.)

81 % of the score are below 63.

7. Given the following statistics, which represents the greatest variability? (1 pt.)

A. $\sigma^2_A = 18$ B. $\sigma^2_B = 56$ C. $\sigma^2_C = 124$ D. $\sigma^2_D = 206$

D represents the greatest variability.

8. Compute the variance and standard deviation for the following scores: (9 pts.)

2, 2, 4, 4, 7, 15, 15, 23

X	minus 9	squared
2	-7	49
2	-7	49
4	-5	25
4	-5	25
7	-2	4
15	6	36
15	6	36
23	14	196
72	0	420

$$72/8 = 9$$

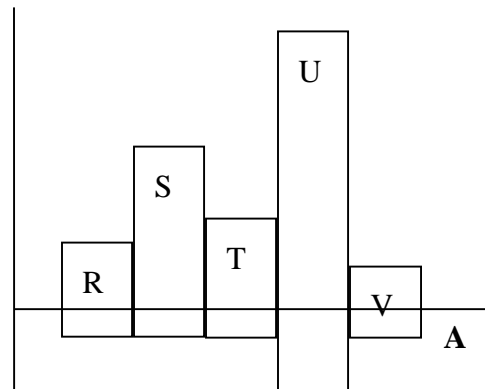
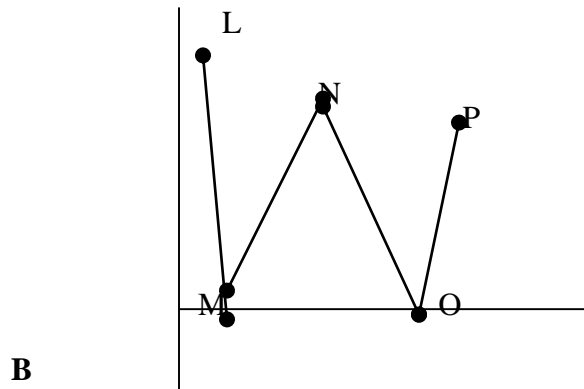
$$\begin{aligned} \text{Variance} \\ &= 420/8 \\ &= \underline{\underline{52.5}} \end{aligned}$$

$$\text{Standard deviation} = \text{square root of } 52.5 = \underline{\underline{7.24}}$$

Interpret the above standard deviation: (2 pts.)

From the calculation above, the standard deviation tell us that we can get 2/3rd of all the score.

Use the following frequency distributions to answer questions 9-13.



9. These two frequency distributions are named, respectively:

A. Frequency Polygon (1 pt.)

B. Histogram (1 pt.)

10. In frequency distribution **A**, the score points have been labeled. Which one represents the mode? (1 pt.)

“U” represents the mode.

11. In frequency distribution **B**, which bar represents the highest score? (1 pt.)

Highest score is “O”

12. According to the convention discussed in class, which one of the following data sets would be appropriate to display in a frequency distribution like **B**? (1 pt.)

Both distribution would be appropriate.

- A. How many people bought cherries, apples, peas, grapes, or peaches?
- B. How many people owned 0 cars, 1 car, 2 cars, 3 cars, or 4 cars?
- C. How many people weighted 140, 150, 160, 170, and 180 pounds?
- D. How many people lived in Dayton, Kettering, Fairborn, Xenia, or Beavercreek?

13. According to the convention discussed in class, discrete data would be plotted using distribution(s) _____. (1 pt.)

Distribution "B" which is Histogram.

The famous water conservationist, Mrs. Sippy, obtained the following results from her research study. Use this information to help answer questions 14-19.

$$\begin{array}{ll} \bar{X} = 71 & \sigma = 14 \\ Q_3 = 76 & \text{Mode} = 67 \\ Q_2 = 60 & \text{Range} = 63 \\ Q_1 = 56 & N = 300 \end{array}$$

14. What type of distribution is this (symmetrical, positive skew, negative skew)? (1 pt.)

$$\begin{array}{l} Q1-Q2 < Q1-Q3 \\ 56-60 < 56-76 \\ -4 < -20 \end{array}$$

This is a positive skewness

15. What evidence is there from the above results to support your decision in #14? (1 pt.)

Evidence: $Q1-Q2 < Q1-Q3$

$\bar{X} >$ than the median ($Q2$)

Here we report the median because it is less affected.

16. Approximately 200 people scored between ___ and ___ (1 pt.)

$$71+14 = 85$$

$$71-14 = 57 \quad \text{between } \underline{57} \text{ and } \underline{85}$$

17. How many people obtained scores below 56? (1 pt.)

$$25\% \text{ of the people} = 25\% \text{ of } 300 = \underline{75 \text{ people}}$$

18. The most frequently occurring score is (1 pt.)

$$\underline{\text{Mode} = 67}$$

19. What is the variance of this distribution? (1 pt.)

$$\text{Since } \sigma = 14$$

$$\text{Variance} = 14^2 = \underline{196}$$

Two physical education researchers, X. R. Sizemore and Ben Dover, obtained the following stem and leaf plots of test scores. **Use these to help answer questions 20-26.**

<u>Plot W</u>			<u>Plot X</u>		
Stem	leaf	f	Stem	leaf	f
5	0000	4	5	0	1
4		0	4	89	2
4		0	4	77	2
4		0	4	55	2
4	233	3	4	223	3
4	011	3	4	001	3
3	89	2	3	88	2
3	777	3	3	66	2
3	445	3	3	45	2
3	33	2	3	2	1

<u>Plot Y</u>			<u>Plot Z</u>		
Stem	leaf	f	Stem	leaf	f
5	0	1	5	000	3
4	89	2	4		0
4	6677	4	4	66666	5
4	555	3	4		0
4	222	3	4	45	2
4	000	3	4	23	2
3	8	1	3		0
3		0	3	66666	5
3		0	3		0
3	223	3	3	222	3

20. **By inspection**, (DON'T CALCULATE!) which distribution(s) would be described as symmetrical? (1 pt.)

All of the above

21. What is the reason for your choice in #20? (1 pt.)

Their mean=median

22. **By inspection**, (DON'T CALCULATE!) what would be the most appropriate measure of central tendency to describe distribution Y: (1 pt.)

It is symmetrical

23. What is the reason for your choice in #22? (1 pt.)

The mean and the median have nearly equal values. Here we report the mean because it uses all the scores.

24. What interval size is being used? (1 pt.)

Interval of 5

25. What is the median for distribution Z? (1 pt.)

Median = 45

26. What is the mode for distribution W? (1 pt.)

Mode = 50

The following test scores were obtained by our history teacher, Cybil Warr.

Use these to help answer questions 27-29.

32, 30, 21, 16, 23, 16, 34, 24, 18, 18, 22, 35, 16, 22, 40, 19, 24

27. Arrange these scores in a stem and leaf plot. (10 pts.)

16, 16, 16, 18, 18, 19, 21, 22, 22, 23, 24, 24, 30, 32, 34, 35, 40.

Range: Highest = 40, Lowest = 16

$40 - 16 = 24 + 1 = 25$

$25 / 5 = 5$ range

Intervals	Stem	Leaf	<i>f</i>
40-44	4	0	1
35-39	3	5	1
30-34	3	0,2,4	3
25-29	2	0	0
21-24	2	1,2,2,3,4,4	6
16-20	0	16,16,16,19	4
			15

28. Find the mean, median, and mode for these scores.

16, 16, 16, 18, 18, 19, 21, 22, 22, 23, 24, 24, 30, 32, 34, 35, 40.

Total of raw scores = 410

$$410/17 = 24.12$$

$$\text{Mean: } 410/17 = \underline{24.12} \quad (1 \text{ pt.})$$

$$\text{Median: } 22+22 = 44/2 = \underline{22} \quad (1 \text{ pt.})$$

$$\text{Mode: } \underline{16} \quad (1 \text{ pt.})$$

29. Describe the distribution of these scores as to its symmetry or skewness. (1 pt.)

The distribution of the scores is symmetrical

Give STATISTICAL evidence for your choice. (1 pt.)

In symmetrical distribution the mean and the media are equally the same. Therefore we report the mean. It is the only one that that uses all the scores.

The well-known dental hygienist, Hallie Tosis, has obtained the attached print out for a set of test scores. **Use these to help answer questions 30-36.**

Univariate Procedure

Variable = SCORE

		Moments	
N	54	Sum Wgts	54
Mean	30.72222	Sum	1659
Std Dev	4.70414	Variance	22.12893
Skewness	-1.3026	Kurtosis	0.801235
USS	52141	CBS	1172.833
CV	15.31185	Std Mean	0.640152
T:Mean=0	47.99205	Pr>=IMI	0.0001

Num ^ = 0	54	Num > 0	54
M(Sign)	27	Pr>= M	0.0001
Sgn Rank	742.5	Pr>= S	0.0001
W:Normal	0.8143	Pr<W	0.0001

100% Max	35	99%	35
75% Q3	34	95%	35
50% Med	33	90%	35
25% Q1	26	10%	23
0% Min	17	5%	21
		1%	17

Range 18
Q3-Q1 5
Mode 34

Stem	Leaf	#	
3	44444444444444445555555555	22	+-----+
3	222333333	9	*-----*
3	00011111	8	+
2	8899	4	
2	67	2	+-----+
2	445	3	
2	23	2	
2	11	2	0
1	9	1	0
1	7	1	0

30. Describe the symmetry or skewness of this distribution. (1 pt.)

It is a negative skewness

31. What number from the printout supports your decision in #30? (1 pt.)

-1.3026

32. What number from the printout tells you this is NOT a bell curve? (1 pt.)

0.0001

33. In the boxplot there is a horizontal broken line that is the bottom of the box. What number from the print out does this represent? (1 pt.)

26 (Q₁)

34. The middle 50% of the scores in this distribution fall between what two values? (Note: Give the numbers from the printout.) (1 pt.)

It is between Q₁ and Q₃ = 26 and 34

35. What is P₁₀? Give the number from the printout. (1 pt.)

10 % percent quartile = 23

36. How many outliers are there? (1 pt.)

There are 4 outliers (17, 19, 21 and 21)