

Show all necessary steps to receive full credit. Any incorrect statement will be penalized.

4 1. Classify each as nominal level, ordinal level, interval level, or ratio level of measurement.

a. Rating of movies as G, PG, and R

nominal

b. Number of candy bars sold on a fund drive

ratio

c. Classification of automobiles as subcompact, compact, standard, and luxury

ordinal

d. Weights of suitcases on a commercial airliner

ratio

4 2. Identify which type of sampling is used: random, systematic, stratified, or cluster.

a. Use the postal ZIP Codes to divide the state into regions. Pick a random sample of 10 Zip Code to divide the state into regions. Pick a random sample of 10 Zip Codes areas and then include all the businesses in each selected Zip Code area.

cluster

b. Group the businesses according to type: medical, shipping, retail, manufacturing, financial, construction, restaurant, hotel, tourism, other. Then select a random sample of 10 businesses from each businesses type.

stratified

c. A researcher divided subjects into two groups according to gender and then selected a few members from each group for her sample.

stratified

d. First, number all the businesses in the Island Business Directory. Second, select a starting place at random, and then use every 50<sup>th</sup> business listed until you have 100 businesses.

systematic

3 3. If the average number of textbooks in professor's office is 16, the standard deviation is 5, and the average age of the professors is 43, with a standard deviation of 8. Which data set is more variable?

$$\text{Textbook} : \text{CVar} = \frac{S}{\bar{X}} \cdot 100\% = \frac{5}{16} \cdot 100\% = 31.25\%$$

$$\text{Age} : \text{CVar} = \frac{S}{\bar{X}} \cdot 100\% = \frac{8}{43} \cdot 100\% = 18.60\%$$

The number of textbooks in professor's office is more variable.

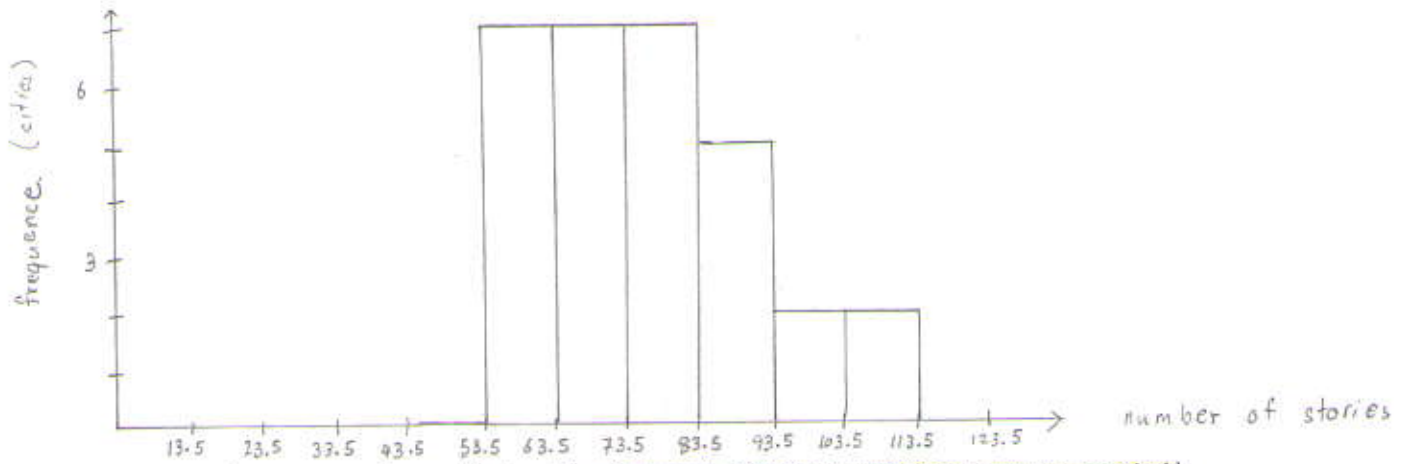
3 4. Please help TTC to calculate his average payoff (*weighted mean*) from three investments' payoff: AAPL \$10000, BAC \$3000, NOK \$1000. The percent of investment in each stock are 30%, 50%, 20% respectively.

$$\text{average payoff} = \frac{(10000)(0.3) + (3000)(0.5) + (1000)(0.2)}{0.3 + 0.5 + 0.2}$$

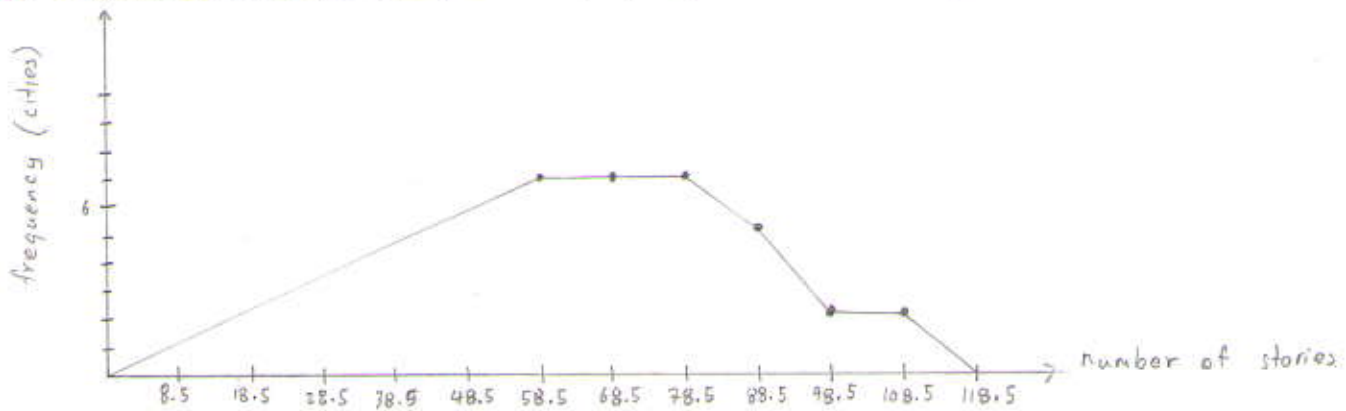
$$= 4700$$

$$\frac{\sum (x \cdot w)}{\sum w}$$

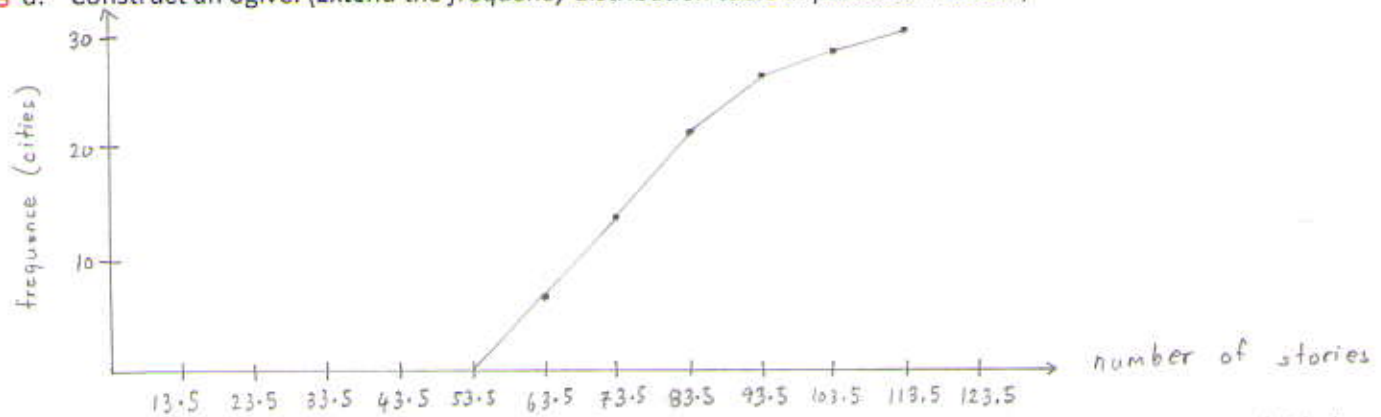
3 b. Construct a frequency histogram. (Extend the frequency distribution table in part a as needed.)



3 c. Construct a frequency polygon. (Extend the frequency distribution table in part a as needed.)



3 d. Construct an ogive. (Extend the frequency distribution table in part a as needed.)



3 e. Find mean. (Write down which formula you used.) Can use raw data or frequency distribution to find mean.

$$\bar{X} = \frac{\sum f \cdot X_m}{n}$$

$$\approx 76.5$$

$$\bar{X} = \frac{\sum X}{n}$$

$$\approx 75.9$$

3 f. Find median. (Use raw data)

$$\text{median} = \frac{75 + 75}{2} = 75$$

- 3 5. The average delivery charge for a refrigerator is \$32. The standard deviation is \$4. Find the minimum percentage data values that will fall in the range of \$20 to \$44. (hint: The problem didn't mention about the kind of distribution, so which theorem will be more appropriate --- Chebyshev's or Empirical?)

$$\bar{x} = 32$$

$$s = 4$$

$$44 = 32 + k(4)$$

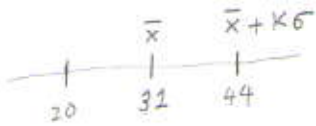
$$12 = 4k$$

$$3 = k$$

$$\left(1 - \frac{1}{k^2}\right) \cdot 100\%$$

$$= \left(1 - \frac{1}{3^2}\right) \cdot 100\%$$

$$= 88.89\%$$



88.89% of data values will fall in the range of \$20 to \$44.

- 3 6. A student scored 76 on a general science test where the class mean and standard deviation were 82 and 8, respectively; he also scored 53 on a psychology test where the class mean and standard deviation were 58 and 3, respectively. In which class was his relative position higher?

$$\text{Science: } z = \frac{x - \bar{x}}{s}$$

$$= \frac{76 - 82}{8}$$

$$= -0.75$$

$$\text{psychology: } z = \frac{x - \bar{x}}{s}$$

$$= \frac{53 - 58}{3}$$

$$\approx -1.67$$

The student did better on Science test.

7. The number of stories in each of the world's 30 tallest buildings follows.

54 55 55 55 56 60 60 (64) 65 69

70 70 71 72 (75) (75) 77 78 79 80

80 85 (88) 88 88 90 100 102 105 110

- 5 a. Construct a grouped frequency distribution. (hint: use the formula to find the ideal number of classes.)

$$\text{range} = 110 - 54$$

$$= 56$$

$$\# \text{ of classes} = 1 + 3.3 \log(30)$$

$$\approx 6$$

$$\text{class width} = \frac{56}{6}$$

$$\approx 10$$

# of stories	f	boundaries	$X_M$	c.f	$X_M^2$	$f \cdot X_M$	$f \cdot X_M^2$
54 - 63	7	53.5 - 63.5	58.5	7	3422.25	409.5	23955.75
64 - 73	7	63.5 - 73.5	68.5	14	4692.25	479.5	32845.75
74 - 83	7	73.5 - 83.5	78.5	21	6162.25	549.5	43135.75
84 - 93	5	83.5 - 93.5	88.5	26	7832.25	442.5	39161.25
94 - 103	2	93.5 - 103.5	98.5	28	9702.25	197	19404.50
104 - 113	2	103.5 - 113.5	108.5	30	11772.25	217	23544.50
						2295	182047.5

a
b, c, d

2 g. Find mode.

$$\text{mode} = \{55, 88\}$$

7 h. Find standard deviation and variance. (Write down which formula you used.)

$$s = \sqrt{\frac{n \sum f \cdot x_m^2 - (\sum f \cdot x_m)^2}{n(n-1)}} \quad s^2 = \frac{n \sum f \cdot x_m^2 - (\sum f \cdot x_m)^2}{n(n-1)}$$

$$\approx 14.9 \quad \approx 223.4$$

raw data

$$s = 15.7$$

$$s^2 = 246.6$$

2 i. Find the z-score of 77.

$$z = \frac{x - \bar{x}}{s} = \frac{77 - 75.9}{15.7} \approx 0.07 \quad \leftarrow \text{raw data}$$

$$z = \frac{77 - 76.5}{14.9} \approx 0.03 \quad \leftarrow \text{frequency distribution.}$$

2 j. Find  $Q_1$  and  $Q_3$ .

$$Q_1 = 64, \quad Q_3 = 88$$

4 k. Find all the outliers. Show your work.

$$\begin{aligned} \text{IQR} &= Q_3 - Q_1 & Q_1 - 1.5(\text{IQR}) & & Q_3 + 1.5(\text{IQR}) \\ &= 88 - 64 & = 64 - 1.5(24) & & = 88 + 1.5(24) \\ &= 24 & = 28 & & = 124 \end{aligned}$$

no outliers.

3 l. Find the value corresponds to the 40<sup>th</sup> percentile.

$$L = \frac{n \cdot p}{100} = \frac{(30)(40)}{100} = 12$$

Since locator is whole number, we need to take the average of 12<sup>th</sup> and 13<sup>th</sup> data value.

$$P_{40} = \frac{70 + 71}{2} = 70.5$$

3 m. Find the percentile rank of the data 90. Interpret the meaning of result percentile.

$$\begin{aligned} \% \text{ rank} &= \frac{25 + 0.5}{30} \cdot 100 \\ &= 85\% \end{aligned}$$

85% of data values fall below the value 90.

+5 n. List 5-data summary and construct boxplot. Then discuss the skewness of the data.

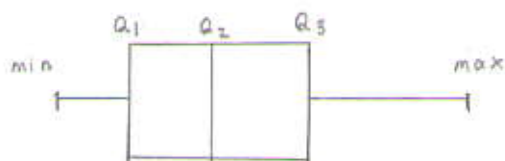
$$\text{min} = 54$$

$$Q_1 = 64$$

$$Q_2 = 75$$

$$Q_3 = 88$$

$$\text{max} = 110$$



Positively skewed //