

Practice Exercises in Descriptive Statistics

- Classify the following data as nominal, ordinal, interval or ratio. Classify each of the data sets as qualitative or quantitative.
 - The length of time it takes each of 15 telephone installers to hook up a wall phone (*ratio, quantitative*)
 - The style of music preferred by each of 30 randomly selected radio listeners (*nominal, qualitative*)
 - The arrival time of the 5 P.M. train from New York to Newark (*interval, quantitative*)
 - A sample of 100 customers in a fast-food restaurant is asked to rate their hamburger on the following scale: poor, fair, good, excellent (*ordinal, qualitative*)
- The state department of education receives the average attendance per day for each of the nine weeks in a grading period from all the state's counties, plus the information of full attendance in each county. Is this an example of raw data? Why or why not? (*No, because it has been analyzed to some extent to get averages for each week's attendance and to get percentages of full attendance. Raw data would be the actual number of absences for each day or week of the time period.*)
- The number of raisins in each of 14 miniboxes (1/2-oz size) was counted for a generic brand and for Sunmaid brand raisins. The data are shown here:
Generic brand: 25, 26, 26, 26, 26, 28, 27, 26, 25, 28, 24, 28, 27, 25
Sunmaid: 25, 28, 25, 28, 29, 24, 28, 24, 24, 28, 30, 24, 22, 27
 - What are the mean and standard deviation for the generic brand?
 - What are the mean and standard deviation for the Sunmaid brand?
 - Compare the centers and variabilities of the two brands using the results of parts **a** and **b**.
 - Find the median, the upper and lower quartiles and the IQR for each of the two data sets. *Generic: $\tilde{X} = 26$, $Q_1 = 25$, $Q_3 = 27.25$, $IQR = 2.25$; Sunmaid: $\tilde{X} = 26$, $Q_1 = 24$, $Q_3 = 28$, $IQR = 4$*
 - Construct two box plots on the same horizontal scale to compare the two sets of data.
 - If we can assume that none of the boxes of raisins are being underfilled (that is, they all weigh approximately $\frac{1}{2}$ ounce), what do your results say about the average number of raisins for the two brands? (*The average is nearly the same; individual raisin sizes are more variable for Sunmaid raisins.*)
- There are 4 private secondary schools in the city. The schools are all similar in terms of tuition costs and academic curriculum. During the past year, the average enrolment capacity per school was 420 children. This year, by adding faculty and facilities, two of the schools have increased their capacity by 10%, which will then make all four schools' enrolment capacities equal to each other. What capacities did the 4 schools have before the increase? *Ans: 440*
- The number of television viewing hours per household and the prime viewing times are two factors that affect television-advertising income. A random sample of 25 households in a particular viewing area produced the following estimates of viewing hours per household: 3.0, 6.5, 5.0, 7.5, 9.0, 6.0, 8.0, 12.0, 5.0, 2.0, 7.5, 4.0, 1.0, 10.0, 6.5, 15.0, 5.5, 3.5, 8.0, 1.0, 12.0, 6.0, 3.0, 3.5, 5.0

- a. Calculate the sample mean and the sample standard deviation.
 $\bar{x} = 6.22, s = 3.497$
- b. Find the percentage of the viewing hours per household that falls into the interval $\bar{x} \pm 2s$. Compare with the corresponding percentage given by the empirical rule. What may be deduced as to the distribution of the data? *96%*

6. The lengths of time (in months) between the onset of a particular illness and its recurrence were recorded: 2.1, 9.0, 14.7, 19.2, 4.1, 7.4, 14.1, 8.7, 1.6, 3.7, 4.4, 2.0, 9.6, 6.9, 18.4, 0.2, 1.0, 24.0, 3.5, 12.6, 2.7, 6.6, 16.7, 4.3, 0.2, 8.3, 2.4, 1.4, 11.4, 23.1, 32.3, 3.9, 7.4, 3.3, 6.1, 0.3, 2.4, 8.2, 18.0, 5.6, 9.9, 1.6, 8.2, 1.2, 13.5, 1.3, 18.0, 5.8, 26.7, 0.4

- a. Examine the data and count the number of observations that fall into the intervals $\bar{x} \pm s$, $\bar{x} \pm 2s$, $\bar{x} \pm 3s$, *37, 47, 49*
- b. Do the percentages that fall into these intervals agree with Chebyshev's theorem? With the Empirical rule? *Yes, no*
- c. Why might the Empirical rule be unsuitable for describing these data? *Distribution is skewed right*

7. The Wall Street Journal Stock Market Data Bank reports the number of shares traded on the New York Stock Exchange in half-hourly intervals. Following are the combined numbers of shares traded (in millions of shares) at half-hourly intervals for three recent days

Shares traded in millions	No. of half-hourly periods
5 – 9	8
10 – 14	17
15 – 19	8
20 – 24	3
25 – 29	2
30 – 34	1

- a. Find the mean and standard deviation of this distribution. *14.05, 6.04*
- b. Find the median and the mode. *12.9, 12*
- c. Calculate the Pearsonian coefficient of skewness for the distribution of shares traded on the NYSE at half-hourly intervals. *+0.55*

8. Given the distribution of the clerical aptitude scores as follows:

Score	Frequency
10 – 19	1
20 – 29	6
30 – 39	9
40 – 49	31
50 – 59	42
60 – 69	32
70 – 72	17
80 – 89	10
90 – 99	2

Find the first quartile, ninth decile and the fifteenth percentile. *46.4, 77.7, 41.6*