

Ch. 3.1 # 3, 4, 7, 30, 31, 32

Find (a) the mean, (b) the median, (c) the mode, and (d) the midrange.

3) **High Temperatures** The reported high temperatures (in degrees Fahrenheit) for selected world cities on an October day are shown below. Which measure of central tendency do you think best describes these data?

62	72	66	79	83	61	62	85	72	64	74	71
42	38	91	66	77	90	74	63	64	68	42	

(a) Mean

$$\text{Mean} = \frac{\sum(x)}{n} = \frac{1566}{23} = 68.1$$

(b) Median

38, 42, 42, ..., **68**, ..., 85, 90, 91 Median=68

(c) Mode= 42, 62, 64, 66, 72, 74

(d) Midrange

$$\text{Midrange} = \frac{38+91}{2} = 64.5$$

For the best measure of average, answers will vary.

4) **Observers in the Frogwatch Program** The number of observers in the Frogwatch USA program (a wildlife conservation program dedicated to helping conserve frogs and toads) for the top 10 states with the most observers is 484, 483, 422, 396, 378, 352, 338, 331, 318, and 302. The top 10 states with the most active watchers list these numbers of visits: 634, 464, 406, 267, 219, 194, 191, 150, 130, and 114. Compare the measures of central tendency for these two groups of data.

For Observers:

a. $\bar{X} = \frac{\sum X}{n} = \frac{3804}{10} = 380.4$ b. MD: $\frac{352+378}{2} = 365$ c. Mode: no mode

d. MR = $\frac{484+302}{2} = 393$

For Visits:

$\bar{X} = \frac{\sum X}{n} = \frac{2769}{10} = 276.9$ MD : $\frac{194+219}{2} = 206.5$

Mode: no mode MR = $\frac{114+634}{2} = 374$

The values are higher for observers.

7) **Earthquake Strengths** twelve major earthquakes had Richter magnitudes shown here.

5.4 5.4 6.2 6.2 6.4 6.4 6.5 7.0 7.2 7.2 7.7 8.0

$$\text{a) Mean} = \frac{\sum(x)}{n} = \frac{79.6}{12} = 6.63$$

$$\text{b) Median} = \frac{6.4 + 6.5}{2} = 6.45$$

c) Mode = none

$$\text{d) Midrange} = \frac{5.4 + 8.0}{2} = 6.7$$

Since almost all data is around the mean value (6.63), the mean is a better choice for the average of measurement.

30) Final Grade An instructor grades exams, 20%; term paper, 30%; final exam, 50%. A student had grades of 83, 72, and 90, respectively, for exams, term paper, and final exam. Find the student's final average. Use the weighted mean.

$$\bar{X} = \frac{\sum w \cdot X}{\sum w} = \frac{20(83) + 30(72) + 50(90)}{100} = 83.2$$

31) Final Grade Another instructor gives four 1-hour exams and one final exam, which counts as two 1-hour exams. Find a student's grade if she received 62, 83, 97, and 90 on the 1-hour exams and 82 on the final exam.

$$\bar{X} = \frac{\sum w \cdot X}{\sum w} = \frac{1(62) + 1(83) + 1(97) + 1(90) + 2(82)}{6} = \frac{496}{6} = 82.7$$

32) For these situations, state which measure of central tendency—mean, median, or mode—should be used.

a. The most typical case is desired. Mode

b. The distribution is open-ended. Median

c. There is an extreme value in the data set. Median

d. The data are categorical. Mode

e. Further statistical computations will be needed. Mean

f. The values are to be divided into two approximately equal groups, one group containing the larger values and one containing the smaller values. Median

Section 3-2 # 7, 9, 11, 13, 30, 31, 33, 34, 35, 37, 41, 42

For exercises 7-13, find the range, variance, and standard deviation. Assume the data represent samples, and use the shortcut formula for the unbiased estimator to compute the variance and standard deviation.

$$\text{mean} = \bar{X} = \frac{\sum(X)}{n};$$

$$\text{variance} = s^2 = \frac{\sum(X - \bar{X})}{n-1} \text{ or } s^2 = \frac{\sum X^2 - [(\sum X)^2 / n]}{n-1}$$

$$\text{standard deviation} = s = \sqrt{\text{variance}}$$

7) **Police Calls in Schools** the number of incidents where police were needed for a sample of 10 schools in Allegheny County is 7, 37, 3, 8, 48, 11, 6, 0, 10, 3. Are the data consistent or do they vary? Explain your answer.

	x	x^2
	0	0
	3	9
	3	9
	6	36
	7	49
	8	64
	10	100
	11	121
	37	1369
	48	2304
Sum	133	4061

$$\text{Mean} = \frac{\sum(X)}{n} = \frac{133}{10} = 13.3$$

$$\text{Range} = 48 - 0 = 48 \quad n=10$$

$$\text{Variance} = s^2 = \frac{\sum X^2 - [(\sum X)^2 / n]}{n-1}$$

$$= \frac{4061 - (133)^2 / 10}{9}$$

$$= \frac{4061 - 1768.9}{9} = \frac{2292.1}{9} = 254.7$$

$$\text{Standard deviation} = \sqrt{254.7} \approx 16$$

The data vary widely.

9) **Precipitation and High Temperatures** the normal daily high temperatures (in degrees Fahrenheit) in January for 10 selected cities are as follows.

50, 37, 29, 54, 30, 61, 47, 38, 34, 61

The normal monthly precipitation (in inches) for these same 10 cities is listed here.

4.8, 2.6, 1.5, 1.8, 1.8, 3.3, 5.1, 1.1, 1.8, 2.5

	x	x^2
	29	841
	30	900
	34	1156
	37	1369
	38	1444
	47	2209
	50	2500
	54	2916

Which set is more variable?

	61	3721
	61	3721
Sum	441	20,777

For the Temperature:

$$\text{Mean} = \frac{\sum(X)}{n} = \frac{441}{10} = 44.1 \quad \text{Range} = 61 - 29 = 32 \quad n=10$$

$$\begin{aligned} \text{Variance} = s^2 &= \frac{\sum X^2 - [(\sum X)^2 / n]}{n-1} = \frac{20,777 - [(441)^2 / 10]}{9} \\ &= \frac{20,777 - 19,448.1}{9} = \frac{1328.9}{9} = 147.6 \end{aligned}$$

$$\text{Standard deviation} = \sqrt{147.6} = 12.15$$

$$\text{CVar} = \frac{12.15}{44.1} = 27.6\%$$

For the Precipitation:

	x	x^2
	4.8	23.04
	2.6	6.76
	1.5	2.25
	1.8	3.24
	1.8	3.24
	3.3	10.89
	5.1	26.01
	1.1	1.21
	1.8	3.24
	2.5	6.25
Sum	26.3	86.13

$$= \frac{86.13 - 69.17}{9} = \frac{16.96}{9} = 1.88$$

$$\text{Standard deviation} = \sqrt{1.88} = 1.37$$

$$\text{CVar} = \frac{1.373}{2.63} = 52.2\%$$

$$\text{Mean} = \frac{\sum(X)}{n} = \frac{26.3}{10} = 2.63$$

$$\text{Range} = 5.1 - 1.1 = 4.0 \quad n=10$$

Variance =

$$s^2 = \frac{\sum X^2 - [(\sum X)^2 / n]}{n-1} = \frac{86.13 - [(26.3)^2 / 10]}{9}$$

Precipitation is more variable since CVar is higher.

11) **Stories in the Tallest Buildings** The number of stories in the 13 tallest buildings for two different cities is listed below. Which set of data is more variable?

Houston: 75, 71, 64, 56, 53, 55, 47, 55, 52, 50, 50, 50, 47
 Pittsburgh: 64, 54, 40, 32, 46, 44, 42, 41, 40, 40, 34, 32, 30

Houston:

$$\bar{X} = \frac{725}{13} = 55.8$$

$$R = 75 - 47 = 28$$

$$s^2 = \frac{n\sum X^2 - (\sum X)^2}{n(n-1)} = \frac{13(41,379) - 725^2}{13(13-1)} = \frac{12,302}{13(12)} = 78.859$$

$$s = \sqrt{78.859} = 8.88$$

Pittsburgh:

$$\bar{X} = \frac{539}{13} = 41.5$$

$$R = 64 - 30 = 34$$

$$s^2 = \frac{n\sum X^2 - (\sum X)^2}{n(n-1)} = \frac{13(23,413) - 539^2}{13(13-1)} = \frac{13,848}{13(12)} = 88.769$$

$$s = \sqrt{88.769} = 9.42$$

The data for Pittsburgh seems more variable. Using the coefficient of variation to compare the data:

For Houston, $CVar = \frac{8.88}{55.8} = 15.91\%$. For Pittsburgh, $CVar = \frac{9.42}{41.5} = 22.7\%$

The data for Pittsburgh is more variable using the coefficients of variation.

13) **Hardcover Bestsellers** the number of weeks on The New York Times Best Sellers list for hardcover fiction is

1 4 2 2 3 18 5 5 10 4 3 6 2 2 22

Use the range rule of thumb to estimate the standard deviation. Compare the estimate to the actual standard deviation.

$$s = \frac{\text{range}}{4} = \frac{22-1}{4} \approx 5.25,$$

$$s^2 = \frac{\sum X^2 - [(\sum X)^2 / n]}{n-1} = \frac{1061 - \left[\frac{(89)^2}{15}\right]}{15-1} \approx 38.1$$

$$s = \sqrt{38.1} \approx 6.2$$

	x	x^2
	1	1
	4	16
	2	4
	2	4
	3	9
	18	324
	5	25
	5	25
	10	100
	4	16
	3	9
	6	36
	2	4
	2	4
	22	484
Sum	89	1061

By the rule of thumb the estimation of the standard deviation is $s=5.25$, and by actually calculating it the standard deviation is $s = 6.2$. Thus, we can say that the estimate is close to the actual standard deviation.

30) **Exam Scores** the average score on an English final examination was 85, with a standard deviation of 5; the average score on a history final exam was 110, with a standard deviation of 8. Which class was more variable?

The coefficients of variation are:

$$\text{English Final:} \quad \text{CVar} = \frac{s}{\bar{X}} = \frac{5}{85} \cdot 100\% = 5.9\%$$

$$\text{History Final:} \quad \text{CVar} = \frac{s}{\bar{X}} = \frac{8}{110} \times 100\% = 7.3\%$$

The history class was more variable.

31) **Ages of accountants** the average age of the accountants at Three Rivers Corp. is 26 years, with a standard deviation of 6 years; the average salary of the accountants is \$31,000, with a standard deviation of \$4000. Compare the variations of age and income.

$$\text{Age:} \quad \text{CVar} = \frac{s}{\bar{X}} = \frac{6}{26} \cdot 100\% = 23.1\%$$

$$\text{Salary:} \quad \text{CVar} = \frac{s}{\bar{X}} = \frac{4000}{31000} \cdot 100\% = 12.9\%$$

Age is more variable.

33) The mean of a distribution is 20 and the standard deviation is 2. Use Chebyshev's theorem.

a) At least what percentage of the values will fall between 10 and 30?

$$k = 5$$

$$\frac{1}{k^2} = \frac{1}{5^2} = \frac{1}{25} = 0.04$$

$$1 - \frac{1}{k^2} = 1 - 0.04 = 0.96 = 96\%$$

b) At least what percentage of the values will fall between 12 and 28?

$$k = 4$$

$$\frac{1}{k^2} = \frac{1}{4^2} = \frac{1}{16} = 0.0625$$

$$1 - \frac{1}{k^2} = 1 - 0.0625 = 0.9375 = 93.75\%$$

34) In a distribution of 160 values with a mean of 72, at least 120 fall within the interval 67–77. Approximately what percentage of values should fall in the interval 62–82? Use Chebyshev's theorem.

$$\frac{120}{160} = 0.75 = 75\% ;$$

$$1 - \frac{1}{k^2} = 0.75 \rightarrow 1 - 0.75 = \frac{1}{k^2} \rightarrow 0.25 = \frac{1}{k^2} \rightarrow k^2 = \frac{1}{0.25} \rightarrow k^2 = 4 \rightarrow k = \sqrt{4} \rightarrow k = 2$$

$$72 + 2s = 77 \rightarrow s = 2.5$$

$$72 + 2.5k = 82 \rightarrow k = 4 \rightarrow 1 - \frac{1}{4^2} = 0.9375 \text{ or } 93.75\%$$

At least 93.75% of the data values will fall in the interval 62–82.

35) **Calories** The average number of calories in a regular size bagel is 240. If the standard deviation is 38 calories, find the range in which at least 75% of the data will lie. Use Chebyshev's theorem.

$$1 - \frac{1}{k^2} = 0.75 \rightarrow 1 - 0.75 = \frac{1}{k^2} \rightarrow 0.25 = \frac{1}{k^2} \rightarrow k^2 = \frac{1}{0.25} \rightarrow k^2 = 4 \rightarrow k = \sqrt{4} \rightarrow k = 2$$

At least 75% of the data values will fall within two standard deviations of the mean; hence, $(240 - 2(38), 240 + 2(38)) = (164, 316)$

At least 75% of the data values will fall between 164 and 316 calories.

37) **Solid Waste Production** The average college student produces 640 pounds of solid waste each year including 500 disposable cups and 320 pounds of paper. If the standard deviation is approximately 85 pounds, within what weight limits will at least 88.89% of all students' garbage lie?

$$1 - \frac{1}{k^2} = 0.8889 \rightarrow 1 - 0.8889 = \frac{1}{k^2} \rightarrow 0.1111 = \frac{1}{k^2} \rightarrow k^2 = \frac{1}{0.1111} \rightarrow k^2 = 9.009 \rightarrow k = \sqrt{9.009} \rightarrow k \approx 3$$

$$(640 - 3(85), 640 + 3(85)) = (385, 895)$$

Therefore at least 88.89% of the data values will fall between 385 and 895 pounds.

41) **Citrus Fruit Consumption** the average U.S. yearly per capita consumption of citrus fruit is 26.8 pounds. Suppose that the distribution of fruit amounts consumed is bell-shaped with a standard deviation equal to 4.2 pounds. What percentage of Americans would you expect to consume more than 31 pounds of citrus fruit per year?

By the Empirical Rule, 68% of consumption is within 1 standard deviation of the mean. Then $\frac{1}{2}$ of 32%, or 16% of consumption would be more than 31 pounds of citrus fruit per year.

42) **Work Hours for College Faculty** the average full-time faculty member in a post-secondary degree-granting institution works an average of 53 hours per week.

- a) If we assume the standard deviation is 2.8 hours, what percentage of faculty members work more than 58.6 hours a week?

$$k = 2$$

$$\frac{1}{k^2} = \frac{1}{2^2} = \frac{1}{4} = 0.25$$

$$\frac{0.25}{2} = 0.125 = 12.5\%$$

12.5% of faculty members work more than 58.6 hours a week.

- b) If we assume a bell-shaped distribution, what percentage of faculty members work more than 58.6 hours a week?

By the Empirical Rule, 95% of members working is within 2 standard deviations of the mean. Then $\frac{1}{2}$ of 5% or

2.5% of faculty member work more than 58.6 hours a week.

Ch 3-3 #11, 13, 14, 22, 23, 24, 25, 27

11) **Exam Scores** A final examination for a psychology course has a mean of 84 and a standard deviation of 4. Find the corresponding Z score for each raw score.

$$\text{a) } z = \frac{x - \mu}{\sigma} = \frac{87 - 84}{4} = 0.75$$

$$\text{b) } z = \frac{x - \mu}{\sigma} = \frac{79 - 84}{4} = -1.25$$

$$\text{c) } z = \frac{x - \mu}{\sigma} = \frac{93 - 84}{4} = 2.25$$

$$\text{d) } z = \frac{x - \mu}{\sigma} = \frac{76 - 84}{4} = -2$$

$$\text{e) } z = \frac{x - \mu}{\sigma} = \frac{82 - 84}{4} = -0.5$$

13) Which has a better relative position: a score of 75 on a statistics test with a mean of 60 and a standard deviation of 10 or a score of 36 on an accounting test with a mean of 30 and a variance of 16?

a. A score of 75 on an exam with $\bar{X} = 60$ and $s = 10$.

$$\frac{x - \bar{x}}{s} = \frac{75 - 60}{10} = 1.5$$

b. A score of 36 on an exam with $\bar{X} = 30$ and $s = 4$

$$\frac{x - \bar{x}}{s} = \frac{36 - 30}{4} = 1.5$$

Neither. The scores have the same relative position.

14) **Test Scores** A student scores 60 on a mathematics test that has a mean of 54 and a standard deviation of 3, and she scores 80 on a history test with a mean of 75 and a standard deviation of 2. On which test did she perform better?

$$z = \frac{x - \mu}{\sigma} = \frac{60 - 54}{3} = 2.0$$

$$z = \frac{x - \mu}{\sigma} = \frac{80 - 75}{2} = 2.5$$

The student performed better in the history test.

22) **Average Weekly Earnings** The average weekly earnings in dollars for various industries are listed below. Find the percentile rank of each value.

804 736 659 489 777 623 597 524 228

Reorder: 228, 489, 524, 597, 623, 659, 736, 777, 804

$$\text{Percentile} = \frac{(\text{number of values below } x) + 0.5}{\text{total number of values}} \cdot 100\%$$

$$\text{For 228, } \frac{0 + .5}{9} \cdot 100\% = 6^{\text{th}} \text{ percentile}$$

$$\text{For 489, } \frac{1 + .5}{9} \cdot 100\% = 17^{\text{th}} \text{ percentile}$$

$$\text{For 524, } \frac{2 + .5}{9} \cdot 100\% = 28^{\text{th}} \text{ percentile}$$

$$\text{For 597, } \frac{3 + .5}{9} \cdot 100\% = 39^{\text{th}} \text{ percentile}$$

$$\text{For 623, } \frac{4 + .5}{9} \cdot 100\% = 50^{\text{th}} \text{ percentile}$$

$$\text{For 659, } \frac{5 + .5}{9} \cdot 100\% = 61^{\text{st}} \text{ percentile}$$

$$\text{For 736, } \frac{6 + .5}{9} \cdot 100\% = 72^{\text{nd}} \text{ percentile}$$

$$\text{For 777, } \frac{7 + .5}{9} \cdot 100\% = 83^{\text{rd}} \text{ percentile}$$

$$\text{For 804, } \frac{8 + .5}{9} \cdot 100\% = 94^{\text{th}} \text{ percentile}$$

23) In exercise 22, what value corresponds to the 40th percentile?

What is 40% of n? = 0.40 x 9 = 3.6, round up to 4.

Therefore, the number located in the 4th place which is 597.

24) **Test scores** find the percentile rank for each test score in the data set.

12 28 35 42 47 49 50

$$\text{Percentile} = \frac{(\text{number of values below } x) + 0.5}{\text{total number of values}} \cdot 100\%$$

$$12: \text{ percentile} = \frac{0 + 0.5}{7} \cdot 100\% \approx 7^{\text{th}}$$

$$28: \text{percentile} = \frac{1+0.5}{7} \bullet 100 \approx 21 \text{ st}$$

$$35: \text{percentile} = \frac{2+0.5}{7} \bullet 100 \% \approx 36 \text{ th}$$

$$42: \text{percentile} = \frac{3+0.5}{7} \bullet 100 \% \approx 50 \text{ th}$$

$$47: \text{percentile} = \frac{4+0.5}{7} \bullet 100 \% \approx 64 \text{ th}$$

$$49: \text{percentile} = \frac{5+0.5}{7} \bullet 100 \% \approx 79 \text{ th}$$

$$50: \text{percentile} = \frac{6+0.5}{7} \bullet 100 \% \approx 93 \text{ rd}$$

25) In exercise 24, what value corresponds to the 60th percentile?

What is 60% of n?

$0.6 \times 7 = 4.2$, round it up to 5. Therefore the number in the 5th location which is 47 corresponds to the 60th percentile.

27) What value in Exercise 26 corresponds to the 40th percentile?

Values in Exercise 26: 1.1 1.7 1.9 2.1 2.2 2.5 3.3 6.2 6.8 20.3

What is 40% of n?

$0.40 \times 10 = 4$. Thus, the value halfway between the 4th and 5th place is 2.15 which corresponds to the 40% percentile.

3.4 # 1, 3, 7, 9, 11, 13, 15

Identify the five-number summary and find the interquartile range.

1) 8, 12, 32, 6, 27, 19, 54

Reorder: 6, 8, 12, 19, 27, 32, 54

Low = 6 $Q_1 = 8$ Median = $Q_2 = 19$ $Q_3 = 32$ High = 54

$$IQR = Q_3 - Q_1 = 32 - 8 = 24$$

3) 188 192 316 362 437 589

Low = 188 $Q_1 = 192$ Median = 339 $Q_3 = 437$ High = 589

$$IQR = Q_3 - Q_1 = 437 - 192 = 245$$

Use each box plot to identify the maximum value, minimum value, median, first quartile, third quartile, and interquartile range.

7) Low = 3, High = 11, median = 8, $Q_1 = 5$, $Q_3 = 9$, and $IQR = Q_3 - Q_1 = 4$

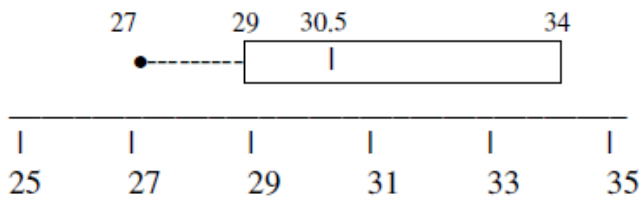
9) Low = 200, High = 325, median = 275, $Q_1 = 225$, $Q_3 = 300$, and $IQR = Q_3 - Q_1 = 75$

11) **Earned Run Average—Number of Games Pitched** Construct a boxplot for the following data and comment on the shape of the distribution representing the number of games pitched by major league baseball's earned run average (ERA) leaders for the past few years.

30 34 29 30 34 29 31 33 34 27
30 27 34 32

Reorder: 27, 27, 29, 29, 30, 30, 30, 31, 32, 33, 34, 34, 34, 34

Low = 27, $Q_1 = 29$, median = $\frac{30+31}{2} = 30.5$, $Q_3 = 34$ High = 34

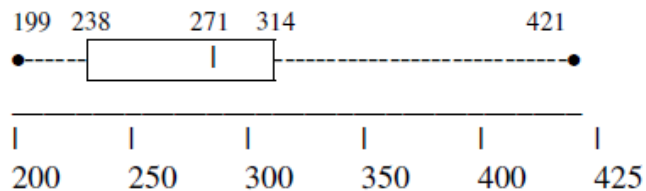


13) **State Sites for Frogwatch** Construct a boxplot for these numbers of state sites for Frogwatch USA. Is the distribution symmetric?

421 395 314 294 289 253 242 238 235 199

Reorder: 199 235 238 242 253 289 294 314 395 421

$MD = \frac{253+289}{2} = 271$
 $Q_1 = 238$ $Q_3 = 314$



No, the distribution is not symmetric.

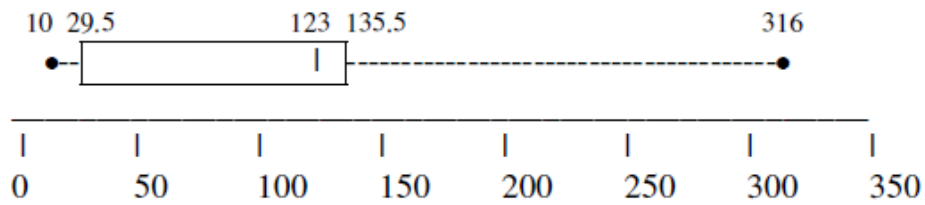
15) **Tornadoes in 2005** Construct a boxplot and comment on its skewness for the number of tornadoes recorded each month in 2005.

33 10 62 132 123 316 138 123 133 18

Reorder: 10 18 33 62 123 123 132 133 138 316

$$MD = \frac{123+123}{2} = 123$$

$$Q_1 = 29.5 \quad Q_3 = 135.5$$



Based on the position of the median, the distribution is negatively skewed. Based on the lengths of the lines, the distribution is positively skewed.