

Chapter 2

- 2.2 The mean profit per partner is \$642,500 calculated using the “=AVERAGE()” command in EXCEL or summing the profits and dividing by the number of firms: $12,850/20 = 642.5$
- 2.4 The total amount of contributions of those teachers was \$4.114 million(=\$242×17000).
- 2.6 a. The average workforce of the women-owned companies is 1.85(= 12 million/6.5 million).
 b. The average workforce of the Fortune 500 companies is 23400(=11,700,000/500).
- 2.8 If a distribution is unimodal, continuous and symmetric, then the mean, median and mode coincide.
- 2.10 a. The mean is 67.78 ($= \frac{50+35+40+65+80+90+95+85+70}{9}$)
 b. The median is 70.
 c. Any value can be a mode because each value occurs only once.
- 2.12 a. Using the ASCII data file EX2-12.PRN, the mean(weighted average) and the median are calculated to be 16.47 percent and 16.5 percent respectively as in the following table.

x_i	f_i	$x_i f_i$	Cumulative f_i
19.0	250	4750	250
18.5	425	7862.5	675
17.0	1200	20400	1875
16.5	235	3877.5	2110<-- Median = 16.5
16.0	110	1760	2220
15.5	750	11625	2970
15.0	935	14025	3905
Sum	3905	64300	

$$\text{Mean} = 64300/3905 = 16.47$$

- b. The distribution is left skewed because the mode(17%) > median(16.5%) > mean(16.47%)

2.14 The medians for three years are approximately 15, 14.375 and 22.5 weeks respectively, which determined as follows:

Median for FY 89 = 15

Median for FY 90 = $10 + \frac{13.5-10}{4}(15-10) = 14.375$

Median for FY 91 = $20 + \frac{9.5-8}{3}(25-20) = 22.5$

Class	m_i	f_i	Cumulative f_i
0-5	2.5	5	5
5-10	7.5	1	6
10-15	12.5	9	15<-- Median class for FY89
15-20	17.5	7	22
20-25	22.5	3	25
25-30	27.5	4	29
30+	?	0	29
Sum		29	
0-5	2.5	2	2
5-10	7.5	8	10
10-15	12.5	4	14<-- Median class for FY90
15-20	17.5	8	22
20-25	22.5	3	25
25-30	27.5	0	25
30+	?	1	26
Sum		26	
0-5	2.5	4	4
5-10	7.5	1	5
10-15	12.5	1	6
15-20	17.5	2	8
20-25	22.5	3	11<-- Median class for FY91
25-30	27.5	3	14
30+	?	4	18
Sum		18	

Although the medians can be approximated by assuming the observations are evenly distributed within the median classes, the means cannot be calculated because the highest class is "30+," making its midpoint undefined.

2.16 Using the ASCII data file EX2-16.PRN (read into EXCEL fixed width columns) the medians for the three spreadsheets are calculated to be 7.15, 5.825 and 6.34 respectively, based upon the following data, which are rearranged in the ascending order. The median is better than the mean as a measure of central tendency here because the values in these data sets are qualitative. Each value simply identifies a consumer's degree of satisfaction at each check point and thus, those values cannot be added together in a meaningful way as is done in the mean calculation. Their relative magnitudes have no meaning on a number line.

	CA- SUPERCALC	LOTUS 1-2-3	MICROSOFT EXCEL	

	6.43	5.25	5.10	
	6.48	5.33	5.71	
	6.59	5.41	5.93	
	6.69	5.55	5.95	
	6.78	5.60	6.14	
	6.81	5.71	6.15	
	6.90	5.75	6.17	
	6.96	5.75	6.22	
	7.00	5.79	6.26)), Median range
	7.30	5.86	6.42))-
	7.30	6.19	6.68	
	7.31	6.23	6.77	
	7.34	6.50	6.78	
	7.45	6.64	6.98	
	7.50	6.70	7.00	
	7.50	6.71	7.00	
	7.52	6.78	7.11	
	7.70	7.18	7.53	
Mean	7.09	6.05	6.44	
Median	7.15	5.825	6.34	

2.18 We suspect that the doctor is using the "average survival" to refer to the median because he is comparing an average survival of 7.3 months with his own result that "half" of the patients have survived 23 months or longer.

2.20 Sample mean = 52.09(= 520.9/10): sum of values divided by n

Sample standard deviation = 32.5438(= $\sqrt{1059.10}$: square root of sample variance).

Sample variance = 1059.10(= 9531.9189/9: sum of squared deviations divided by degrees of freedom(n-1)).

Standard error of the mean = 10.29125 = 32.5438/3.16 (the sample standard deviation divided by the square root of the sample size). It is a measure of variability in the distribution of the sample mean.

The other advertisers bought \$1,069 million of commercial time.(= \$1,590 m. - \$520.9 m.)

2.22 As shown in computer printout below, the mean of data set B is 100 times larger (i.e., 10^2 larger) than the mean of data set A. Likewise B's standard deviation is 100 times greater than A's. The variance, however, because it is the square of the standard deviation, is a multiple of 10,000 times larger (i.e., 10^4 larger) for data set B than for data set A. These relationships would hold even if these were population data. But the sum of squared deviations would be divided by 4 for the calculation of the population variance while the sum of squared deviations is divided by 3 for the sample variance.

Variable	Mean	Std. Dev.	Skew.	Minimum	Maximum	Cases
A	.25000	4.3493	-.752	-6.000	4.000	4
B	25.000	434.93	-.752	-600.0	400.0	4

2.24 The approximate sample mean is calculated to be \$9,520 [= 0.1(9900) + 0.2(9700) + 0.4(9500) + 0.3(9300) = 9520]. The approximate sample variance is 36,326.53 (=1780000/49); thus, the approximate sample standard deviation is \$190.595.

m_i	f_i	f_i/n	$(f_i/n)m_i$	$(m_i-9520)^2$	$f_i(m_i-9520)^2$
9900	5	0.1	990	144400	722000
9700	10	0.2	1940	32400	324000
9500	20	0.4	3800	400	8000
9300	15	0.3	2790	48400	726000
sum	50	1.0	9520		1780000

2.26 Without knowing the time required to process those appeals that have lasted more than 30 months, it is difficult to say whether the distributions in 1990 or 1991 are more skewed. With 4 observations above 30 months, 1991 has the potential to be the most skewed.

2.28 As shown in the following table, the CEO compensation(X) tends to deviate from its mean in the same direction as the number of headquartered companies(Y) deviates from its mean. That is, the product of $(x_i - \bar{x})$ and $(y_i - \bar{y})$ tends to be positive. This implies that the CEO compensation(X) and the number of headquartered companies(Y) tend to move in the same direction, thus are positively related to each other.

	x_i	$x_i - \bar{x}$	y_i	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$
	717237	246211.14	49	36.5	8986706.71
	479772	8746.14	2	-10.5	-91834.50
	463009	-8016.86	1	-11.5	92193.86
	439441	-31584.86	16	3.5	-110547.00
	416066	-54959.86	0	-12.5	686998.21
	390933	-80092.86	4	-8.5	680789.29
	390723	-80302.86	23	10.5	-843180.00
	N.A.		5	-7.5	
Sum	3297181		100		9401126.57
Average	471025.857		12.5		

2.30 Twenty-five is two standard deviations below the mean of 55 and 85 is two standard deviations above 55. Because the distribution is bell shaped, about 95 percent of the calls could be handled between 25 and 85 seconds.

2.32 \$140,000 is 1.5 standard deviations above the mean of \$80,000 and \$20,000 is 1.5 standard deviations below \$80,000. Using Chebychev's theorem, "at least" 56 percent(i.e., $1 - 1/1.5^2$) of total students could have been paid between \$20,000 and \$140,000. Thus, "at most" 44 percent of total students could have been paid above \$140,000(up to the maximum of \$400,000).

2.34 The mean delivery time is calculated to be 156 minutes(=2 hours 36 minutes) for all 84 letters. But the exact median and the mode cannot be calculated because individual delivery time is not available.

m_i	f_i	$m_i f_i$	Cumulative f_i
91	12	1092	12
104	12	1248	24
120	12	1440	36
182	12	2184	48 \leftarrow Approximate median = 182
192	12	2304	60
195	12	2340	72
211	12	2532	84
Sum	84	13140	

Mean = $13140/84 = 156.4286$ minutes

2.36 Assuming that the observations are evenly spread within classes and the range of the first class starts from zero, we can calculate percentiles as follows:

For example,

$$5\text{th percentile} = (0+7.8)/2 = 3.90$$

$$15\text{th percentile} = (7.8+10.7)/2 = 9.25$$

m_i	Cumulative Relative Frequency
3.90	0.05 \leftarrow 5th percentile
7.80	0.10
9.25	0.15 \leftarrow 15th percentile
10.70	0.20
11.70	0.25 \leftarrow 25th percentile
12.70	0.30
13.50	0.35 \leftarrow 35th percentile
14.30	0.40
15.35	0.45 \leftarrow 45th percentile
16.40	0.50
17.70	0.55 \leftarrow 55th percentile
19.00	0.60
20.65	0.65 \leftarrow 65th percentile
22.30	0.70
24.90	0.75 \leftarrow 75th percentile
27.50	0.80
33.00	0.85 \leftarrow 85th percentile
38.50	0.90
59.50	0.95 \leftarrow 95th percentile
80.50	1.00

2.38 The mean, standard deviation, and median as determined in EXCEL with the identified commands are:

Am Herit		Val Forge	
19		6	
31		6	
35		8	
41		9	
96		17	
44.400	=AVERAGE(A2:A6)	9.200	=AVERAGE(D2:D6)
29.947	=STDEV(A2:A6)	4.550	=STDEV(D2:D6)
35	=MEDIAN(A2:A6)	8	=MEDIAN(D2:D6)

Both distribution are right skewed, with a range of 77 and 11 points respectively. As volatility (risk) goes up so does the mean (expected) return.

NOTE TO INSTRUCTORS: Because these are rates of return, in finance it is common to calculate the geometric mean of $(1+r)$, which is the $\exp\{\text{mean}[\ln(1+r)]\}$ and the associated standard deviation, which is $\exp\{\text{stdev}[\ln(1+r)]\}$. Below are the EXCEL calculations and commands.

	0.17395		0.05827
	0.27003		0.05827
	0.30010		0.07696
	0.34359		0.08618
	0.67294		0.15700
	0.352	=AVERAGE(1A2:A16)	0.087 =AVERAGE(D12:D16)
	0.190	=STDEV(A12:A16)	0.041 =STDEV(D12:D16)
	0.30010	=MEDIAN(A12:A16)	0.076961 =MEDIAN(D12:D16)
Geometric			
Mean (1+r)	1.4221	=EXP(A17)	1.0913 =EXP(D17)
Standard			
Dev. (1+r)	1.2091	=EXP(A18)	1.0416 =EXP(D8)
Median			
of (1+r)	1.3500	=EXP(A19)	1.0800 =EXP(D19)

The *Washington Post* article (January 29, 1995) claimed American Heritage Mutual Fund had a standard deviation of "a whopping 18.5," and Valley Forge Mutual Fund had a reported standard deviation of "a mere 1.3." As seen above neither of these standard deviations are appropriate. There is no question, however, that as volatility goes up so does the mean return.

2.40 An approximate median is calculated to be 17.6 percent as follows:

$$\text{Median} = 17 + \frac{0.5-0.41}{0.15}(18-17) = 17.6(\%)$$

Class	f_i/n	Cumulative f_i/n
19+	0.36	1.00
18-19	0.08	0.64
17-18	0.15	0.56<-- Median class
16-17	0.12	0.41
12-16	0.27	0.29
12-	0.02	0.02
Sum	1.00	

To calculate the median, we have to assume the observations are evenly spread within the median class.

2.42 Based upon given information, the medians of number of times of play per year can be provided to be 12 times for skiers, 25 times for tennis players and 25 times for golfers respectively. However, the means and modes cannot be determined from the information.

2.44 It would be possible to observe that 54.2 percent are above average if the distribution of the cost of health insurance is left skewed with some(5.7 percent) small values pulling down the mean. Thus, the median would be higher than the mean in this distribution.

2.46 Using the ASCII data file EX2-46.PRN, the mean automobile output in 1992 and 1991 are calculated to be 144,524 and 142,188.6 respectively. The medians are 62,000 and 49,000 respectively. The mean and median are not equal in each year because the distribution of automobile output is not symmetric.

	1992	1991	% change
	647000	674000	-4.0%
	307000	251000	22.3%
	118837	151000	-21.3%
	100000	112174	-10.9%
	64000	47000	36.2%
	60000	34000	76.5%
	58903	47000	25.3%
	43000	51000	-15.7%
	32500	39712	-18.2%
	14000	15000	-6.7%
Sum	1445240	1421886	83.6%

Mean in 1992 = $1445240/10 = 144524$

Mean in 1991 = $1421886/10 = 142188.6$

Median in 1992 = $(64000+60000)/2 = 62000$

Median in 1991 = $(51000+47000)/2 = 49000$

2.48 The percentage change in average can be regarded as a better measure of change than the percentage change in median. The change in median (either 36.2 percent or 76.5 percent) does not adequately represent the situation in percentage change; 36.5 percent or 76.5 percent are the largest values among the values of percentage change.

- 2.50 a. No.
 b. No.
 c. Yes.
 d. Yes, but the mean and median would be determined to be approximate 4,400 fights also.

2.52 For Austin-Los Angeles Continental airline has the median price of \$258. However, for the remaining two cities, the medians cannot be clearly identified because the medians are not well defined in those data sets of (400, 248, 400, 400, 228) and (164, 113, 280, 113, 113).

2.54 Using the ASCII data file EX2-54.PRN, the descriptive sample statistics are calculated as follows:

Mean = 41,914.8
 Median = 37,750
 Mode = 38,000 or 46,000
 Range = 183,000 - 540 = 182,460
 Variance = 1,332,754,040
 Standard deviation = 36,506.904

2.56 a. The means cannot be calculated because the class 11+ is open. Approximate medians are calculated to be 3.46 days and 5.63 days as follows:

Median for those who currently spend

$$= 3 + \frac{0.50-0.44}{0.39}(6-3) = 3.46$$

Median for those who prefer to spend

$$= 3 + \frac{0.50-0.08}{0.48}(6-3) = 5.63$$

Class	m_i	f_i/n	$m_i(f_i/n)$	Cumulative f_i/n
0	0	0.10	0	0.10
1-2	1.5	0.34	0.51	0.44
3-5	4	0.39	1.56	0.83<-- Median class
6-10	8	0.13	1.04	0.96
11+	?	0.04	?	1.00
Sum		1.00		
0	0	0.02	0	0.02
1-2	1.5	0.06	0.09	0.08
3-5	4	0.48	1.92	0.56<-- Median class
6-10	8	0.39	3.12	0.95
11+	?	0.05	?	1.00
Sum		1.00		

b. The above data are not a proper form for a frequency distribution because there is an open class as mentioned above. This implies that it is not possible to calculate the mean with this data set.

c. The approximate median is a better measure of central location because the median can be determined without any trouble caused by the open class.

2.58 The sample mean is calculated to be \$21.618. The sample median is \$18.42. The sample variance is calculated to be \$24.886 and hence the standard deviation is \$4.989. This data cannot be considered to be consistent with the rule of normality because the mean is very different from the median; the distribution is not bell shaped.

	x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
	27.44	5.822	33.8957
	26.69	5.072	25.7252
	18.42	-3.198	10.2272
	17.97	-3.648	13.3079
	17.57	-4.048	16.3863
Sum	108.09		99.5423

Sample mean = $108.09/5 = 21.618$

Sample variance = $99.5423/4 = 24.88557$

2.60 Using the ASCII data file EX2-60.PRN, the following are calculated:

Body Style	Mean
1	1,546.7
2	9,196.4
3	23,592.8
4	33,931.5
5	Not calculated because 1994 missing

The average yearly production has increased with the introduction of new body styles. Although the model average may show a trend in Corvette popularity, if the number of years that each model is produced is too different then model average may not be comparable.

2.62 Using the LOTUS spreadsheet file EX2-62.WK1, the following is calculated:

Average capital/asset ratio = 16.21 percent

The ratio of (average capital)/(average asset)

= $23631.85/161433.5 = 14.64$ percent

They are different because the average capital/asset ratio is the simple arithmetic average of capital/asset ratios of 100 banks while the ratio of (average capital)/(average asset) is nothing but the sum of 100 weighted capital/asset ratios (using the asset/(sum of asset) ratio as each weight) as we did the same in Exercise 2.61.

2.64 Using the LOTUS spreadsheet file EX2-64.WK1 provided on the computer disk, average sales and profit growth rates are calculated for each country as shown in the following table:

	Sales	Profit
U.S. firms:	6.77%	-7.73%
Netherlands/U.K. firms:	4.69%	2.94%
Japan firms:	6.43%	-12.22%
Germany firms:	12.75%	20.00%
France firms:	13.00%	6.00%
Italy firm:	16.00%	N.A.
Switzerland firms:	11.00%	26.00%
Australia firm:	16.00%	9.00%

Rank		%	%
1992	Company (Country)	Change in Sales (90-91)	Change in Profit (90-91)
2	Exxon (U.S.)	-1	12
4	Philip Morris (U.S.)	10	11
5	General Electric (U.S.)	3	3
6	Wal-Mart Stores (U.S.)	26	20
7	AT&T (U.S.)	20	-81
8	Merck (U.S.)	12	19
9	IBM (U.S.)	-6	-109
10	Coca-Cola (U.S.)	13	17
14	Bristol-Myers Squibb (U.S.)	8	18
15	Du Pont (U.S.)	-3	-25
19	Procter & Gamble (U.S.)	12	11
20	General Motors (U.S.)	-2	-151
21	Johnson & Johnson (U.S.)	11	15
26	GTE (U.S.)	2	7
27	PepsiCo (U.S.)	10	-1
31	Abbott Laboratories (U.S.)	12	13
33	Mobil (U.S.)	-3	-1
34	Pfizer (U.S.)	8	-10
35	BellSouth (U.S.)	1	-7
36	Amoco Corp (U.S.)	-10	-20

37 Chevron (U.S.)	-4	-40
41 American Home Products (U.S.)	4	12
42 Ford Motor (U.S.)	-10	-277
45 Minnesota Mining & Mfg. (U.S.)	2	-12
53 Bell Atlantic (U.S.)	0	2
54 Eli Lilly (U.S.)	10	17
55 Microsoft (U.S.)	56	66
56 American Int'l Group (U.S.)	19	8
57 Walt Disney (U.S.)	6	-23
60 Southwestern Bell (U.S.)	2	5
64 Hewlett-Packard (U.S.)	10	2
67 Atlantic Richfield (U.S.)	N.A.	N.A. *
69 Ameritech (U.S.)	2	-6
70 Waste Management (U.S.)	25	-15
73 McDonald's Corp (U.S.)	-1	7
75 Pacific Telesis (U.S.)	2	-1
76 Nynex (U.S.)	-3	-1
77 Texaco (U.S.)	-9	-11
78 Kellogg (U.S.)	12	20
81 Anheuser-Busch (U.S.)	2	12
83 Dow Chemical (U.S.)	-5	-32
85 BankAmerica Corp (U.S.)	4	28
86 Schlumberger (U.S.)	16	16
87 U S West (U.S.)	6	-5
89 Federal Natl Mortgage (U.S.)	7	24
90 Home Depot (U.S.)	38	46
96 Sears Roebuck (U.S.)	2	29
98 Boeing (U.S.)	6	13
99 Pacific Gas & Electric (U.S.)	3	4
<u>Average of U.S. Firms</u>	6.77	-7.73
1 Royal Dutch/Shell (Netherlands/U.K.)	-2	-34
12 British Telecom (U.K.)	7	38
13 Glaxo Holdings (U.K.)	7	13
18 Unilever (Netherlands/U.K.)	4	4
44 SmithKline Beecham (U.K.)	4	17
46 B.A.T. Industries (U.K.)	7	21
47 British Petroleum (U.K.)	-1	-76
49 Guinness (U.K.)	17	16
50 British Gas (U.K.)	10	-4
58 Hanson (U.K.)	7	7
61 Grand Metropolitan (U.K.)	-7	4
63 Marks & Spencer (U.K.)	3	2
65 BTR (U.K.)	0	3
74 Imperial Chemical (U.K.)	-3	-6

82 Sainsbury (J.) (U.K.)	13	13
91 Wellcome PLC (U.K.)	9	29
<u>Average of Netherlands/U.K. Firms</u>	4.69	2.94
3 NTT (Japan)	4	-14
11 Toyota Motor (Japan)	7	-2
16 Sumitomo Bank (Japan)	-2	-21
17 Mitsubishi Bank (Japan)	-4	-26
22 Fuji Bank (Japan)	-12	-34
24 Dai-Ichi Kangyo Bank (Japan)	-6	-37
25 Industrial Bank of Japan (Japan)	-9	-22
28 Tokyo Electric Power (Japan)	7	-20
29 Sanwa Bank (Japan)	0	-30
32 Sakura Bank (Japan)	78	19
43 Matsushita Electric (Japan)	10	10
51 Hitachi (Japan)	9	9
52 Nomura Securities (Japan)	-1	-55
66 Kansai Electric Power (Japan)	8	0
71 Seven-Eleven (Japan)	19	27
72 Toshiba Corp. (Japan)	10	-8
79 Bank of Tokyo (Japan)	-2	-38
88 Nippon Steel (Japan)	8	-22
92 Mitsubishi Heavy (Japan)	13	44
93 Long-Term Credit Bank (Japan)	3	-39
94 Ito-Yokado (Japan)	10	18
97 Tokai Bank (Japan)	-9	-12
100 Chubu Electric Power (Japan)	7	-28
<u>Average of Japan Firms</u>	6.43	-12.22
23 Allianz Holding (Germany)	11	N.A. *
38 Daimler-Benz (Germany)	11	11
39 Siemens (Germany)	16	19
48 Deutsche Bank (Germany)	13	30
<u>Average of German Firms</u>	12.75	20.00
62 Assicurazioni Generali (Italy)	16	N.A. *
59 Elf Aquitaine (France)	15	-8
84 Alcatel Alsthom (France)	11	20
<u>Average of French Firms</u>	13.00	6.00
30 Nestle (Switzerland)	9	9
40 Roche Holding (Switzerland)	19	56
80 Sandoz (Switzerland)	9	15
95 Ciba-Geigy (Switzerland)	7	24
<u>Average of Switzerland Firms</u>	11.00	26.00

 * excluding N.A. when calculating the average.

2.66 Using the ASCII data file EX2-66.PRN, the following calculations can be done: (Notice that the ASCII file EX2-66.PRN consists of two columns for the total of 610 observations; the first column is the variable SEX and the second column is the variable SALARY. For SEX, 1 denotes that a person is a female and 0 denotes that a person is a male.)

	Number of observations	Mean Salary(\$)	Median Salary(\$)
All	610	40,225.810	36,500
Women	218	34,641.358	33,150
Men	392	43,331.449	39,850

According to this calculation the women are receiving less than the men. The median salaries are \$33,150 versus \$39,850. If the ten highest paid women each receive \$50,000 raises, nothing changes in the median salary for women.

The mean may be viewed as a better measure of central tendency when assessing the cost of removing salary differences because both means would be exactly the same if a total of \$1,894,439.838(= (218)(43331.449 - 34641.358)) is raised for salaries for those 218 women.

2.68 The men could be lying or the women could be lying. If there is no lying, then it may be just a "distributional effect." For example in a society of two men and three women if the two men had sex with each of the three women then the mean number of partners for men would be three and the mean number of partners for women would be two.

2.70 The approximate first and third quartiles are given by EXCEL as 2.86 and 3.18, but only the 3.18 is a unique quartile value. The median is given as 3.02, but as with the first quartile, the median is not well defined here.

2.24	3.02		Column1	
2.30	3.04		-----	-----
2.30	3.05		Mean	3.0048
2.50	3.05		Standard Error	0.042649248
2.63	3.06		Median	3.02
2.65	3.11		Mode	2.86
2.68	3.12		Standard Deviation	0.301575726
2.75	3.12		Sample Variance	0.090947918
2.78	3.13		Kurtosis	0.431614116
2.78	3.13		Skewness	-0.544862284
2.82	3.14		Range	1.31
2.85	3.17		Minimum	2.24
1st> 2.86	3.18<	3rd	Maximum	3.55
2.86	3.21		Sum	150.24
2.86	3.25		Count	50
2.86	3.27		Largest(13)	3.18
2.87	3.28		Smallest(13)	2.86
2.94	3.34			
2.95	3.36			
2.95	3.36			
2.95	3.41			
3.01	3.42			
3.02	3.50			
3.02	3.52			
3.02	3.55			