3.20 (a) Response: Incidence rates of headaches. Explanatory: Usage of aspartame.

(b) Control: Subjects taking placebos. Treatments: Subjects given aspartame.

(c) It reduces the variability that occurs when different subjects take the treatment and control. If some people are more likely to get headaches than others (due to stress, health, etc...) a crossover design helps to equalize this factor for both treatments.

3.21 Response: Mileage. Treatments: Gasoline additive, Concentrations. Blocks: Make of car.

3.22 Response: Occurrence of stomach irritation. Treatment factors: Dosing schedule, form of administration. Confounding variable: Severity of infection, weight, age.

3.23 (a) Let A=method 1, B=method 2, and C=method 3. Then one CR design is {A,B,B,C,A,A,B,C,B}

(b) One randomized block design is

$$\{\underbrace{A,C,B}_{\text{Person 1}}\ ,\ \underbrace{B,A,C}_{\text{Person 2}}\ ,\ \underbrace{B,C,A}_{\text{Person 3}}\ \}.$$

3.24 (a) Let C=Column sum, R=Row sum. Then one CR design assigns the problems to the 20 students as follows:
{C,R,C,C,R,C,R,R,R,C,R,C,C,R,C,R,R,C,R,C}

(b) Classify the students into fast and slow adders. Then within each group of students, randomize 5 C and 5 R questions to the students. Then one randomized block design is

$$\{\underbrace{C,R,R,C,C,R,C,R,R,C}_{\text{Slow}},\underbrace{R,C,R,C,C,R,R,C,R,C,C}_{\text{Fast}}\}.$$

3.25 (a) Let A=aerobic, S=strengthening, and C=control. Then one CR design is
{A,C,C,S,A,A,C,A,A,C,S,S,S,C,A,C,S,S,C,S,A,S,A,C }

(b) One possible randomized block design is

$$\{\underbrace{A,C,S,S,C,C,A,S,C,A,S,A,}_{\text{65-79}}\underbrace{S,A,C,C,A,A,S,C,A,C,S,S}_{\text{80 and older}}\}.$$

3.26 (a) Let S=Standard method, N=New method. Then one CR design is
{S,N,S,S,N,N,S,N,S,S,N,N,S,S,S,N,S,N,N,S,N,N,S,N,S,N,N,S,S,N}

(b) One randomized block design is

$$\{\underbrace{S,N,S,S,N,N,S,N,S,N}_{\text{Tech 1}},\underbrace{N,S,S,N,N,S,N,N,S,S}_{\text{Tech 2}},\underbrace{S,N,S,N,S,S,N,S,N,N}_{\text{Tech 3}}\}.$$

Chapter 4 Solutions

Solutions for Section 4.1

4.1 (a) Categorical (Ordinal).

(b) Categorical (Nominal)

(c) Numerical (Practically Continuous)

(d) Categorical (Nominal)

4.2 (a) Engine Size: Numerical (Continuous)
Number of Cylinders: Numerical (Discrete)
Size of Car: Categorical (Ordinal)
Type of Transmission: Categorical (Nominal)
Gas Guzzler Tax: Categorical (Nominal)
Dealer Cost: Numerical (Practically Continuous)
Theft Rate Index: Numerical (Practically Continuous)

(b) Engine Size: Ratio scale
Number of Cylinders: Ratio scale
Dealer Cost: Ratio scale

4.3 (a) Enrollment: Numerical (Practically Continuous)
Required Entrance Tests: Categorical (Nominal)
Annual Tuition: Numerical (Practically Continuous)
Fields of Study: Categorical (Nominal)
Selectivity: Categorical (Ordinal)
Percent of Applicants Selected: Numerical (Continuous)

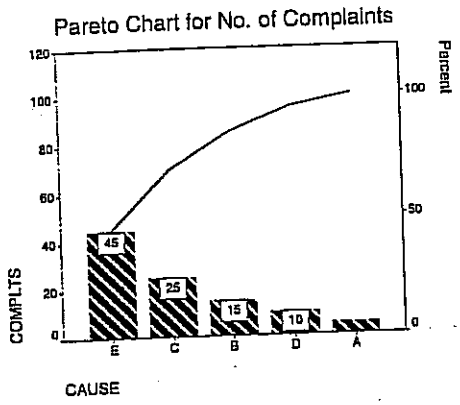
(b) Enrollment: Ratio scale
Annual Tuition: Ratio scale
Percent of Applicants Selected: Ratio scale

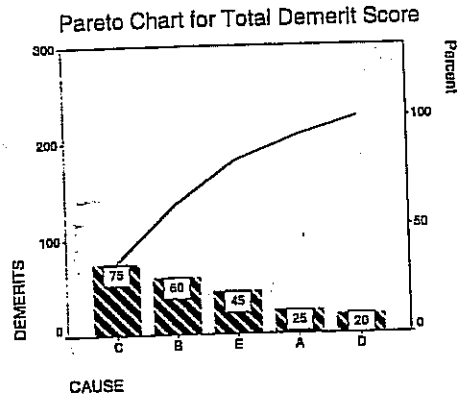
4.4 (a) Sex: Categorical (Nominal)
Age at Graduation: Numerical (Practically Continuous)
Time to Complete Degree: Numerical (Practically Continuous)
Future Plans: Categorical (Nominal)
Citizenship: Categorical (Nominal)
Class Rank: Categorical (Ordinal)
Grade Point Average: Numerical (Continuous)

(b) Age at Graduation: Ratio scale
Time to Complete Degree: Ratio scale
Grade Point Average: Interval scale

Solutions for Section 4.2

4.5

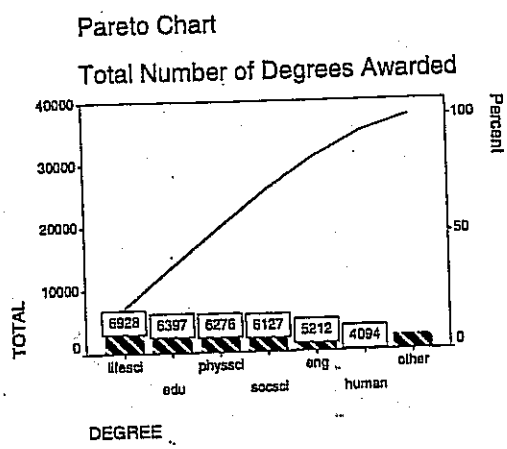




E(45%), C(25%), and B(15%) account for more than 80% of all complaints.

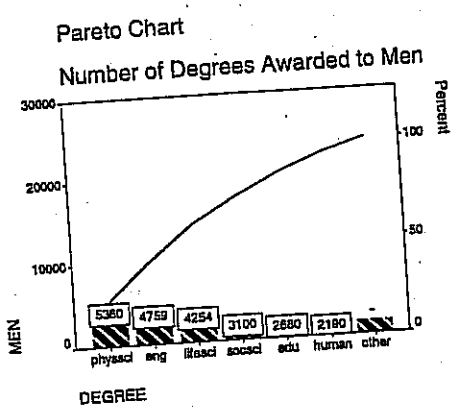
4.6

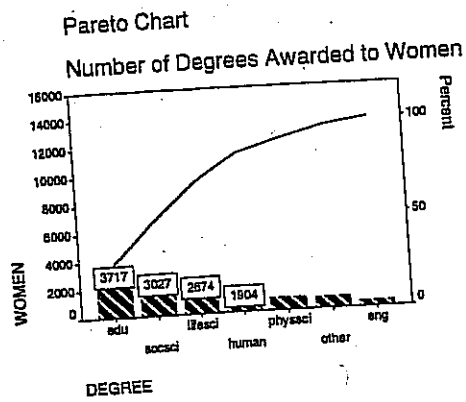
Field	Number of Degrees	Relative Frequency
Physical Sciences	6276	0.168
Engineering	5212	0.139
Life Sciences	6928	0.185
Social Sciences	6127	0.164
Humanities	4094	0.109
Education	6397	0.171
Other	2417	0.065
Total	37451	1.000



The degrees awarded are pretty evenly distributed among the professional fields.

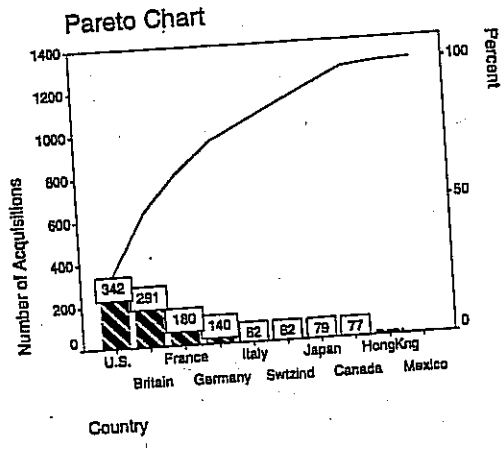
4.7 (a) The number of doctorate degrees awarded to women were: 916 in physical sciences, 453 in engineering, 2674 in life sciences, 3027 in social sciences, 1904 in humanities, 3717 in education, and 836 in other professional fields. The number of doctorate degrees awarded to men were: 5360 in physical sciences, 4759 in engineering, 4254 in life sciences, 3100 in social sciences, 2190 in humanities, 2680 in education, and 1581 in other professional fields.





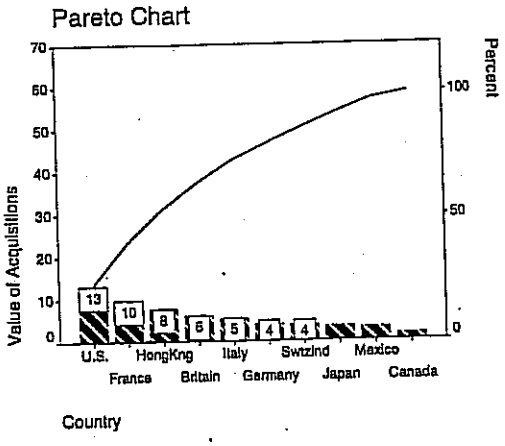
(b) The degrees awarded most to men and women are different. Men are more likely to earn degrees in physical sciences, engineering, and life sciences, while women are more likely to earn degrees in life sciences, social sciences, humanities, and education.

4.8 (a)



The U.S. and Britain have the largest numbers of acquisitions, accounting for almost 50% of all acquisitions.

(b)



The U.S. has the largest acquisition values, followed by France. Together they account for about 40% of all acquisition values.

Solutions for Section 4.3

4.9 (a) The median will be larger because the long left tail will shift the mean left.

(b) The sample skewness will be negative.

(c) One would expect Q_1 to be farther from the median because the long left tail will shift the first quartile farther to the left.

4.10 (a) The distribution is positively skewed.

(b) One would expect Q_3 to be farther from the median because the distribution is positively skewed.

(c)

$$Q_1 = x_{(12.75)} = x_{(12)} + 0.75\left(x_{(13)} - x_{(12)}\right) = 13 + (0.75)(14 - 13) = 13.75.$$

$$\text{Median: } Q_2 = x_{(25.5)} = x_{(25)} + 0.5\left(x_{(26)} - x_{(25)}\right) = 20 + (0.5)(20 - 20) = 20.$$

$$Q_3 = x_{(38.25)} = x_{(38)} + 0.25\left(x_{(39)} - x_{(38)}\right) = 33 + (0.25)(34 - 33) = 33.25.$$

As expected, Q_3 is farther from the median.

(d) The mean will be larger than the median because the positive skew will shift the mean up. Check: $\bar{x} = 25.3 > Q_2 = 20.0$.

(e) The IQR is $33.25 - 13.75 = 19.5$. The lower fence is $13.75 - 1.5 \times 19.5 = -15.5$. The upper fence is $33.25 + 1.5 \times 19.5 = 62.5$. Therefore, 64 and 69 should be outliers.

4.11

(a)

$$\text{Mean: } \overline{x} = \frac{97.0 + 97.2 + ... + 99.8}{30} = 98.563.$$

$$\text{Median: } Q_2 = x_{(15.5)} = x_{(15)} + 0.5\left(x_{(16)} - x_{(15)}\right) = 98.6 + (0.5)(98.6 - 98.6) = 98.6.$$

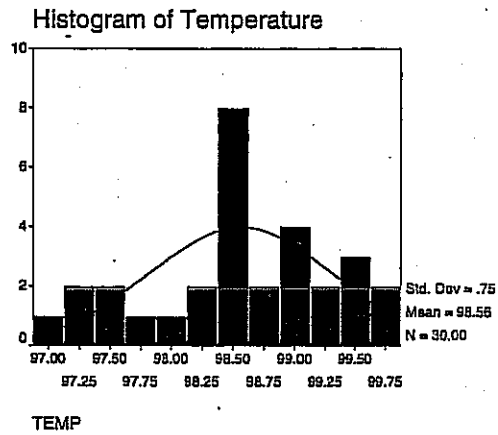
$$\text{SD: } s = \sqrt{\frac{(97.0 - 98.563)^2 + (97.2 - 98.563)^2 + ... + (99.8 - 98.563)^2}{30 - 1}} = 0.751.$$

$$Q_1 = x_{(7.75)} = x_{(7)} + 0.75\left(x_{(8)} - x_{(7)}\right) = 97.9 + (0.75)(98.2 - 97.9) = 98.125.$$

$$Q_3 = x_{(23.25)} = x_{(23)} + 0.25\left(x_{(24)} - x_{(23)}\right) = 99.1 + (0.25)(99.2 - 99.1) = 99.125.$$

(b) There are no outliers, since no observations fall outside the fences ($LF = 98.125 - 1.5 \times 1 = 96.625$ and $UF = 99.125 + 1.5 \times 1 = 100.625$).

(c)



The shape is unimodal with no outliers or skew.

4.12 (a)

$$\text{Min: } 452.$$

$$Q_1 = x_{(12)} = 850.$$

$$\text{Median: } Q_2 = x_{(24)} = 1331.$$

$$Q_3 = x_{(36)} = 1737.$$

$$\text{Max: } 3830.$$

This summary suggests a symmetric distribution, since the quartiles are equidistant from the median. Note, however, that the maximum is much farther from Q_3 than the minimum is from Q_1, possibly because it is an outlier.

(b)

$$\text{Approx.:} \quad s \approx \frac{\text{IQR}}{1.34} = \frac{1737 - 850}{1.34} = 661.94$$

$$\text{Actual:} \quad s = \sqrt{\frac{(1468 - 1369)^2 + (909 - 1369)^2 + ... + (850 - 1369)^2}{47 - 1}} = 693.67.$$

The actual SD is roughly the same as the approx. SD.

(c) A 10% trimmed mean trims off the outside $n\alpha = 47 \times 0.1 = 4.7 = 4$ observations when computing the mean. So

$$\overline{x}^{(0.1)} = \frac{x_{(5)} + x_{(6)} + ... + x_{(43)}}{47 - 2 \times 4} = 1292.64.$$

This is substantially lower than the sample mean of 1369, which suggests that there are some outliers on the high end of the data.

4.13 (a)

RAINFALL Stem-and-Leaf Plot

Frequency	Stem &	Leaf
2.00	0 .	44
17.00	0 .	55567778888999
11.00	1 .	11123333334
11.00	1 .	55577778888
1.00	2 .	0
4.00	2 .	5667
1.00 Extremes		(>=3830)

Stem width: 1000.00
Each leaf: 1 case(s)

The shape is right-skewed and unimodal, with one outlier.

(b)