Chapter 1

1. Method (c) is probably best, with (e) being the second best.

2. In 1936 only upper middle class and rich people had telephones. Almost all voters have telephones today.

3. No, these people must have been prominent to have their obituaries in the Times; as a result they were probably less likely to have died young than a randomly chosen person.

4. Locations (i) and (ii) are clearly inappropriate; location (iii) is probably best.

5. No, unless it believed that whether a person returned the survey was independent of that person's salary; probably a dubious assumption.

6. No, not without additional information as to the percentages of pedestrians that wear light and that wear dark clothing at night.

7. He is assuming that the death rates observed in the parishes mirror that of the entire country.

8. $12{,}246/.02 = 612{,}300$

9. Use them to estimate, for each present age x, the quantity $A(x)$, equal to the average additional lifetime of an individual presently aged x. Use this to calculate the average amount that will be paid out in annuities to such a person and then charge that person $1 + a$ times that latter amount as a premium for the annuity. This will yield an average profit rate of a per annuity.

10. 64 percent, 10 percent, and 48 percent.

Chapter 2

2. $360/r$ degrees.

6. (d) 3.18
(e) 3
(f) 2
(g) $\sqrt{5.39}$

7. (c) 119.14
(d) 44.5
(e) 144.785

8. Not necessarily. Suppose a town consists of n men and m women, and that a is the average of the weights of the men and b is the average of the weights of the women. Then na and mb are, respectively, the sums of the weights of the men and of the women. Hence, the average weight of all members of the town is

$$\frac{na + mb}{n + m} = ap + b(1 - p)$$

where $p = n/(n + m)$ is the fraction of the town members that are men. Thus, in comparing two towns the result would depend not only on the average of the weights of the men and women in the towns but also their sex proportions. For instance, if town A had 10 men with an average weight of 200 and 20 women with an average weight of 120, while town B had 20 men with an average weight of 180 and 10 women with an average weight of 100, then the average weight of an adult in town A is $200\,\frac{1}{3} + 120\,\frac{2}{3} = \frac{440}{3}$ whereas the average for town B is $180\,\frac{2}{3} + 100\,\frac{1}{3} = \frac{460}{3}$.

10. It implies nothing about the median salaries but it does imply that the average of the salaries at company A is greater than the average of the salaries at company B.

11. The sample mean is 110. The sample median is between 100 and 120. Nothing can be said about the sample mode.

12. (a) 40.904
(d) 8, 48, 64

13. (a) 15.808
(b) 4.395

14. Since $\sum x_i = n\bar{x}$ and $(n-1)s^2 = \sum x_i^2 - n\bar{x}^2$, we see that if x and y are the unknown values, then $x + y = 213$ and

$$x^2 + y^2 = 5(104)^2 + 64 - 102^2 - 100^2 - 105^2 = 22{,}715$$

Therefore,

$$x^2 + (213 - x)^2 = 22{,}715$$

Solve this equation for x and then let $y = 213 - x$.

15. No, since the average value for the whole country is a weighted average where the average wage per state should be weighted by the proportion of all workers who reside in that state.

19. (a) 44.8
(b) 70.45

20. 74, 85, 92

21. (a) 84.9167
(b) 928.6288
(c) 57.5, 95.5, 113.5

25. (a) .3496
(b) .35
(c) .1175
(d) no
(e) $3700/55 = 67.3$ percent

26. (b) 3.72067
(c) .14567

28. Not if both sexes are represented. The weights of the women should be approximately normal as should be the weights of the men, but combined data is probably bimodal.

30. Sample correlation coefficient is .4838

31. No, the association of good posture and back pain incidence does not by itself imply that good posture causes back pain. Indeed, although it does not establish the reverse (that back pain results in good posture) this seems a more likely possibility.

32. One possibility is that new immigrants are attracted to higher paying states because of the higher pay.

33. Sample correlation coefficient is .7429

34. If $y_i = a + bx_i$ then $y_i - \bar{y} = b(x_i - \bar{x})$, implying that

$$\frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} = \frac{b}{\sqrt{b^2}} = \frac{b}{|b|}$$

35. If $u_i = a + bx_i$, $v_i = c + dy_i$ then

$$\sum (u_i - \bar{u})(v_i - \bar{v}) = bd \sum (x_i - \bar{x})(y_i - \bar{y})$$

and

$$\sum (u_i - \bar{u})^2 = b^2 \sum (x_i - \bar{x})^2, \quad \sum (v_i - \bar{v})^2 = d^2 \sum (y_i - \bar{y})^2$$

Hence,

$$r_{u,v} = \frac{bd}{|bd|} r_{x,y}$$

36. More likely, taller children tend to be older and that is why they had higher reading scores.

37. Because there is a positive correlation does not mean that one is a cause of the other. There are many other potential factors. For instance, mothers that breast feed might be more likely to be members of higher income families than mothers that do not breast feed.