Yellow

1. (a) iv (b) i (c) vi (d) v

2. First, $\bar{x} = \frac{1}{4}(58 + 37 + 29 + 46) = \left(\frac{1}{4}\right)(170) = 42.5$

Then the sample variance is

$$s^2 = \frac{1}{3}\left[(58-42.5)^2 + (37-42.5)^2 + (29-42.5)^2 + (46-42.5)^2\right]$$

$$= \frac{1}{3}\left[(15.5)^2 + (-5.5)^2 + (-13.5)^2 + (3.5)^2\right]$$

$$= \frac{1}{3}(240.25 + 30.25 + 182.25 + 12.25) = \left(\frac{1}{3}\right)(465) = 155.$$

So, the sample standard deviation is

$$s = \sqrt{155} \doteq 12.45$$

3. (a) iv (b) i (c) iii (d) ii

4. (a) It is symmetric, bimodal

(b) median

(c) range $\approx 21 - 6 = 15$

5-number summary: 6, 8.5, 10.5, 13, 21

IQR $\approx 13 - 8.5 = 4.5$

5. (a) The observation that pet ownership and contentment have an association should not be confused — at least without an experiment to establish as fact — with a causal relationship.

(b) The two variables under consideration — "lives with 2 parents (yes/no)", and "avoids gang membership (yes/no)" — are categorical. Correlation is a concept of statistics reserved for pairs of quantitative variables.

6. Let A = "will take a biology course"

B = "will take a chemistry course"

(a) $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B) = 0.41 + 0.32 - 0.21 = 0.52$

(b) We check whether

$P(A \text{ and } B) = 0.21$ and $P(A) \cdot P(B) = (0.41)(0.32) = 0.1312$.

are equal. Since they are not, A and B are not independent at Red Hill U.

7. Sensitive in this list are the mean, the standard deviation, and the range

8. True statements: (i), (iv) and (vi)

9. We are given these probabilities, when a random message from the week is selected:

$P(\text{marked as spam}) = 0.127$

$P(\text{"free"}) = 0.058$

$P(\text{marked as spam and "free"}) = 0.0455$

(a) $$P(\text{"free"} \mid \text{marked as spam}) = \frac{P(\text{marked as spam and "free"})}{P(\text{marked as spam})}$$

$$= \frac{0.0455}{0.127} \doteq \boxed{0.358}$$

(b) $$P(\text{marked as spam} \mid \text{"free"}) = \frac{P(\text{marked as spam and "free"})}{P(\text{"free"})}$$

$$= \frac{0.0455}{0.058} \doteq \boxed{0.784}$$

10. (a) The study is observational in nature. No conditions are imposed on participants by the researchers.

(b) Is there an association between time spent watching TV and the number of aggressive acts committed?

(c) Explanatory variable (quantitative): time spent watching TV
Response variable (quantitative): number of aggressive acts committed

(d) Many possibilities here. For instance:

- level of parental supervision
- number of activities in which the individual participates as a teen