Quick solutions for old midterm

1. (a) numerical continuous
   (b) numerical discrete
   (c) categorical (ratings are always categorical)
   (d) categorical

2. (a)
$$\bar{x} = \frac{16250}{18}, \quad s = \sqrt{\frac{14869900 - (16250)^2/18}{17}}$$
   (b) $M = (930 + 950)/2, Q1 = 850, Q3 = 980, Min = 650, Max = 1070$
   (c)
```
 6 5
 7 4
 7 6
 8 1
 8 558
 9 03
 9 566888
10 00
10 7
```
   Other acceptable choice for a graph would be a boxplot, or a histogram.
   According to the graph, the center of the distribution is at about 950, the data spreads from 650 to 1070, and the distribution is skewed to the left and unimodal.

3. (a) $P(x = 16) = 1 - 0.1 - 0.3 - 0.25 - 0.1 = .25$
   (b) $\mu = 1(0.1) + 2(0.3) + 4(0.25) + 8(0.1) + 16(0.25)$
   (c) $P(3 < x < 16) = 0.25 + 0.1$

4. Use the empirical rule, which states that approximately 95% of measurements fall within two standard deviations of the mean:
   $[48.5 - 2(6.3); 48.5 + 2(6.3)]$

5. Use normal distribution with $\mu = 4.5, \sigma = 1.1$
   (a)
$$\begin{aligned} P(x \leq 5) &= P((x - 4.5)/1.1 \leq (5 - 4.5)/1.1) \\ &= P(z \leq 0.4545) \\ &= 0.6736 \end{aligned}$$
   (b)
$$\begin{aligned} x_{0.2} &= \mu + z_{0.2}\sigma \\ &= 4.5 + (-0.84)1.1 \end{aligned}$$

6. (a) $P(aggressive) = 150/500$
   (b) $P(aggressive|firstborn) = 75/200$

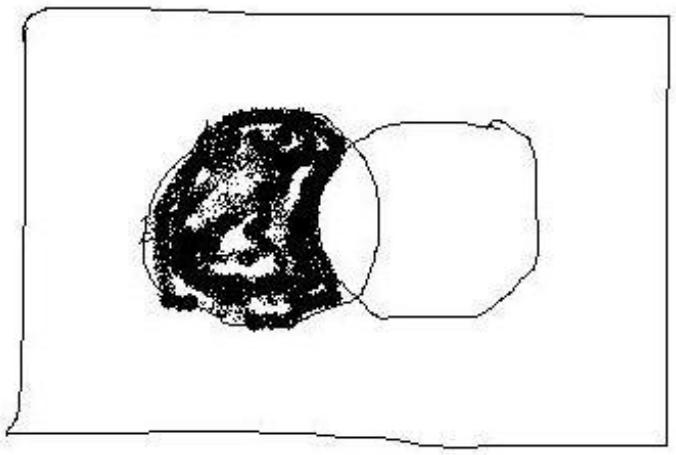
(c) Since $P(aggressive) \neq P(aggressive|firstborn)$, the two events are not independent.

7. (a) The data can be categorical or numerical.

For categorical data one should give a pie- or bar-chart, and a relative frequency table.

For numerical data one should give either the 5-number summary or the mean and standard deviation, and in addition either show a histogram, boxplot, or a stem and leaf plot. Comment.

(b)



(c) Draw a histogram, which shows a heavy skew, or outliers.

(d) A probability distribution describes random variables by giving the possible values, and their probabilities.

(e) Colour in the area under the curve right of $x = 20$