(I) (a)

A: die shows even number: $\{2,4,6\}$

B: die shows number under 3: $\{1,2\}$

$$P(A) = \frac{3}{6} = \frac{1}{2}$$

$$P(B) = \frac{2}{6} = \frac{1}{3}$$

; $A \cap B = \{2\}$
$\Rightarrow P(A \cap B) = \frac{1}{6}$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$= \frac{1}{2} + \frac{1}{3} - \frac{1}{6} = \frac{2}{3}$$

(b) For any 3 sets A, B, C:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

We solve for $P(C)$:

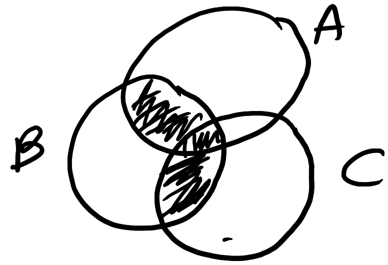
$$P(C) = P(A \cup B \cup C) - P(A) - P(B) + P(A \cap B) + P(A \cap C) + P(B \cap C) - P(A \cap B \cap C)$$

$$= 0.5 - 0.25 - 0.30 + 0.15 + 0.18 + 0.05 - 0.02 = 0.31$$

Note: $P(A \cap B) = 0.15$, $P(B \cap C) = 0.18$

$A_{shaded} = 0.15 + 0.18 - 0.02$
$= 0.31$

$P(B)$ should be $> \underline{0.31}$!



(II) $\bar{X} = \dfrac{\sum x_i f_i}{\sum f_i}$ grouped data

$= \dfrac{1}{100}(60 \cdot 4 + 63 \cdot 8 + \ldots + 84 \cdot 4 + 87 \cdot 2)$

$= 72.45$

$s^2 = \dfrac{\sum (x_i - \bar{x})^2 f_i}{n-1}$

$= \dfrac{1}{99}\left[(60 - \bar{x})^2 \cdot 4 + \ldots + (87 - \bar{x})^2 \cdot 2\right]$

≈ 41.16

$\Rightarrow s \approx 6.41$

$\bar{X} \pm s \approx [66.03, 78.87] = I_1$

$\bar{X} \pm 2s \approx [59.62, 85.28] = I_2$

outside I_1: $\{60, 63, 66, 81, 84, 87\}$

total: $4+8+12+9+4+2=39$

included: $100-39=61\%$

outside I_2: $\{87\}$ $\Rightarrow$ included $100-2=98\%$

Empirical rule says that for normal data I_1 contains 68% and I_2 95%.
Chebyshev's then states that at least $1-\frac{1}{2^2}=\frac{3}{4}$ (or 75%) of data in I_2.

(III) (a) Voluntary response is likely to generate more bias. People who are passionate about the topic would respond. More likely to include more extreme views.

(b) nominal e.g. preferred cereal brand
ordinal e.g. rank ordering 1-100 score for set of movies
interval e.g. set of temperature measurements
ratio e.g. unemployment rate in last 2 months

(c) If $s=\sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}=0$

$\Rightarrow \sum(x_k-\bar{x})^2=0 \Rightarrow (x_k-\bar{x})=0 \Rightarrow x_k=\bar{x}$ for all k

This means all entries of X must be the same.

(d) $w_1 = 0.40;\ w_2 = 0.50,\ w_3 = 0.10$

$$\bar{X}_w = \frac{\sum x_i w_i}{\sum w_i} = \frac{75(0.40) + 62(0.50) + 90(0.10)}{0.40 + 0.50 + 0.10}$$

$$= 70$$

(out of 100 pts since $(\sum w_i) \times 100 = 100$).

(IV) First, sort the data:

$sx = sort(x) =$

$= \{15, 43, 54, 56, 61, 65, 66, 68, 68, 69, 69, 70, 71, 72,$
$77, 78, 79, 85, 87, 88, 89, 98, 99, 99, 100\}.$

15 looks like an outlier

$$L_{25} = \left(\frac{25}{100}\right) n = .25 \times 25 = 6.25 \Rightarrow \lceil L_{25} \rceil = 7$$

$$L_{50} = \left(\frac{50}{100}\right) n = .5 \times 25 = 12.5 \Rightarrow \lceil L_{50} \rceil = 13$$

$$L_{75} = \left(\frac{75}{100}\right) n = .75 \times 25 = 18.75 \Rightarrow \lceil L_{75} \rceil = 19$$

$Q_1 = sx[7] = 66$ $\qquad$ $min = 15$

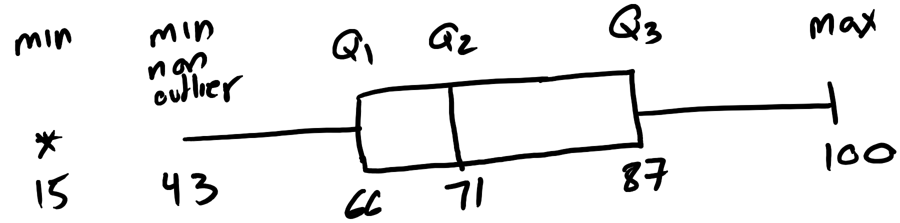
$Q_2 = sx[13] = 71$ $\qquad$ $max = 100$

$Q_3 = sx[19] = 87$ $\qquad$ $IQR = Q_3 - Q_1 = 21$

$Q_1 - 1.5\,IQR = 66 - 1.5 \times 21 = 34.5$

$Q_3 + 1.5\,IQR = 87 + 1.5 \times 21 = 118.5$

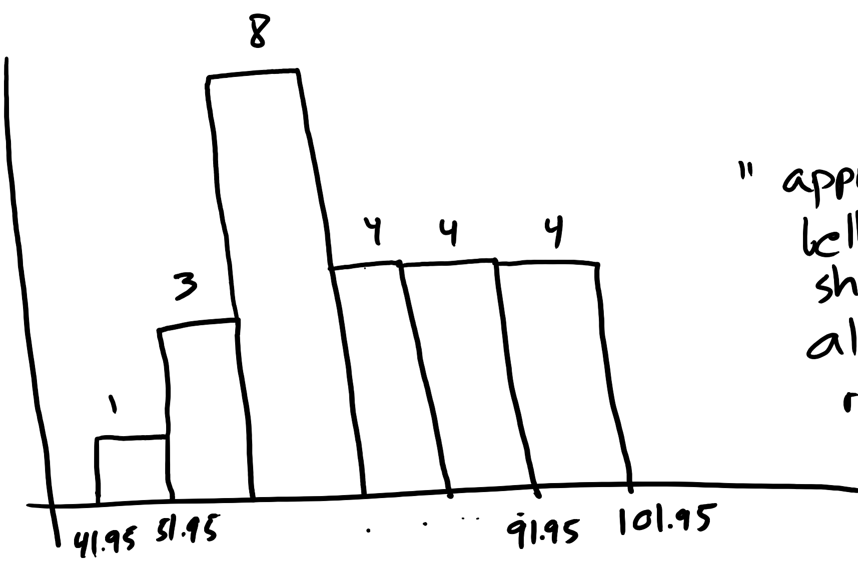
The only value outside $(34.5, 118.5)$ is 15, which is an outlier as suspected.



Histogram (e.g. use class width = 10)

42 - 51.9	1
52 - 61.9	3
62 - 71.9	8
72 - 81.9	4
82 - 91.9	4
92 - 101.9	4

24 samples excluding 15 (outlier)



" approx bell shaped " although only roughly so

(V)

(a) $P((3 \cap C) \cup (6 \cap D))$ mutually exclusive events

$= P(3 \cap C) + P(6 \cap D) = \frac{1}{52} + \frac{1}{52} = \frac{2}{52} = \frac{1}{26}$

(b) $P(H') = 1 - P(H) = 1 - \frac{13}{52} = \frac{3}{4}$

(c) $P(10' \cap S') = P((10 \cup S)') =$

$= 1 - P(10 \cup S)$

$= 1 - [P(10) + P(S) - P(10 \cap S)]$

$= 1 - [\frac{4}{52} + \frac{13}{52} - \frac{1}{52}] = 1 - \frac{4}{13}$

$= \frac{9}{13}$

(d) $x = \{1, 2, 3\}$ $\underline{y = 2x+1}$

$y = \{3, 5, 7\}$

since y and x have perfect <u>linear</u> relationship, it follows that $r = 1$.

(e) We want $P(A \cap B')$.

$(A \cap B') \cup (A \cap B) = A$; $(A \cap B') \cap (A \cap B) = \phi$

$P(A) = P(A \cap B') + P(A \cap B) - \underbrace{P(\phi)}_{=0}$

$\Rightarrow P(A \cap B') = P(A) - P(A \cap B)$

$$\Rightarrow P(A \cap B') = \frac{1}{2} - \frac{1}{6} = \frac{1}{3}$$