S1 Mock (MA)

coffee   empty jar

Q1) $C \sim N(101.80, 0.72^2)$ $E \sim N(260, 5.45^2)$

$P(\text{required}) = \left[P(E < 266)\right] \times \left[P(C < 100)\right]$

$P(E < 266) = P\left(Z < \frac{266-260}{5.45}\right) = P(Z < 1.10)$

$= \boxed{0.8643}$

$P(C < 100) = P\left(Z < \frac{100-101.8}{0.72}\right) = P(Z < -2.5)$

$= P(Z > 2.50) = 1 - P(Z < 2.50)$

$= 1 - 0.9938 = \boxed{0.0062}$

$P(\text{required}) = 0.0062 \times 0.8643 = \boxed{0.0054}$

Q2a) 23. (most common)

b) $\frac{n}{4} = Q_1 = \frac{42}{4} = 10.5 \rightarrow 11^{th}$ observation $\rightarrow \boxed{17 = Q_1}$

$\frac{n}{2} = Q_2 = \frac{42}{2} = 21 \rightarrow$ midpoint of 21^{st} & 22^{nd}

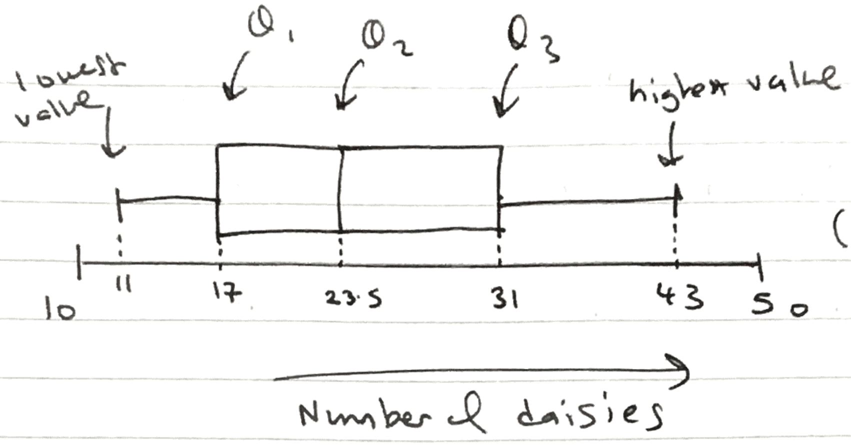
value $= \frac{23+24}{2} = \boxed{23.5}$

= median

$\frac{3n}{4} = Q_3 = \frac{3}{4} \times 42 = 31.5 \rightarrow 32^{nd}$ value $= Q_3$

$= \boxed{31} = Q_3$

c)



(ROUGH).

d) $Q_3 - Q_2 = 31 - 23.5 = 7.5$
$Q_2 - Q_1 = 23.5 - 17 = 6.5$

$7.5 > 6.5$
$\therefore$ there is a positive skew.

e) a 'back-to-back' stem & leaf diagram can be used

Q3a) $\bar{y} = \dfrac{\sum fy}{\sum f} = \dfrac{-467}{200}$ //

$Var(y) = \dfrac{\sum fy^2}{\sum f} - (\bar{y})^2 = \dfrac{9179}{200} - \left(\dfrac{-467}{200}\right)^2$

$= 40.443..$

$\therefore \sigma_y = \sqrt{40.443} = 6.359..$ //

but $y = \dfrac{x - 755}{2.5}$

$\therefore \bar{x} = 2.5\bar{y} + 755 = 749.16$ // $= \boxed{749.2}$

$\therefore \sqrt{Var(x)} = 2.5 \times 6.359.. = \boxed{15.9} = \sigma_x$.

(Variance not affected by $\pm$).

b) s.d = 15.9 < $\frac{2}{3}$ IQR (almost half the IQR)

the s.d shows less variation than the IQR for the lifetimes of bulbs so use s.d.

Q4a) P(correct 3rd time) = $0.4 \times 0.4 \times 0.6 = \boxed{0.096}$

b)

a	1	2	3	4
P(A=a)	0.6	0.24	0.096	0.064

c) $P(\text{required}) = \underbrace{P(A=1) + P(A=2) + P(A=3)}_{\text{if } A=1,2 \text{ or } 3 \text{ then this means the PIN is correctly entered in less than 4 attempts}} + \overbrace{(0.4^3)(0.6)}^{P(\text{success on } 4^{th} \text{ try})}$

$= 0.6 + 0.24 + 0.096 + 0.064 + 0.4^3(0.6$

$= \boxed{0.974}$

d) $E(A) = \sum aP(A=a) = 0.6 + 2(0.24) + 3(0.096) + 4(0.0384)$

$= 1.624$

$E(A^2) = \sum a^2P(A=a) = 0.6 + 4(0.24) + 9(0.096) + 16(0.0384)$

$= 3.448$

$Var(A) = 3.448 - (1.624)^2 = E(A^2) - [E(A)]^2$

$= \boxed{0.811}$

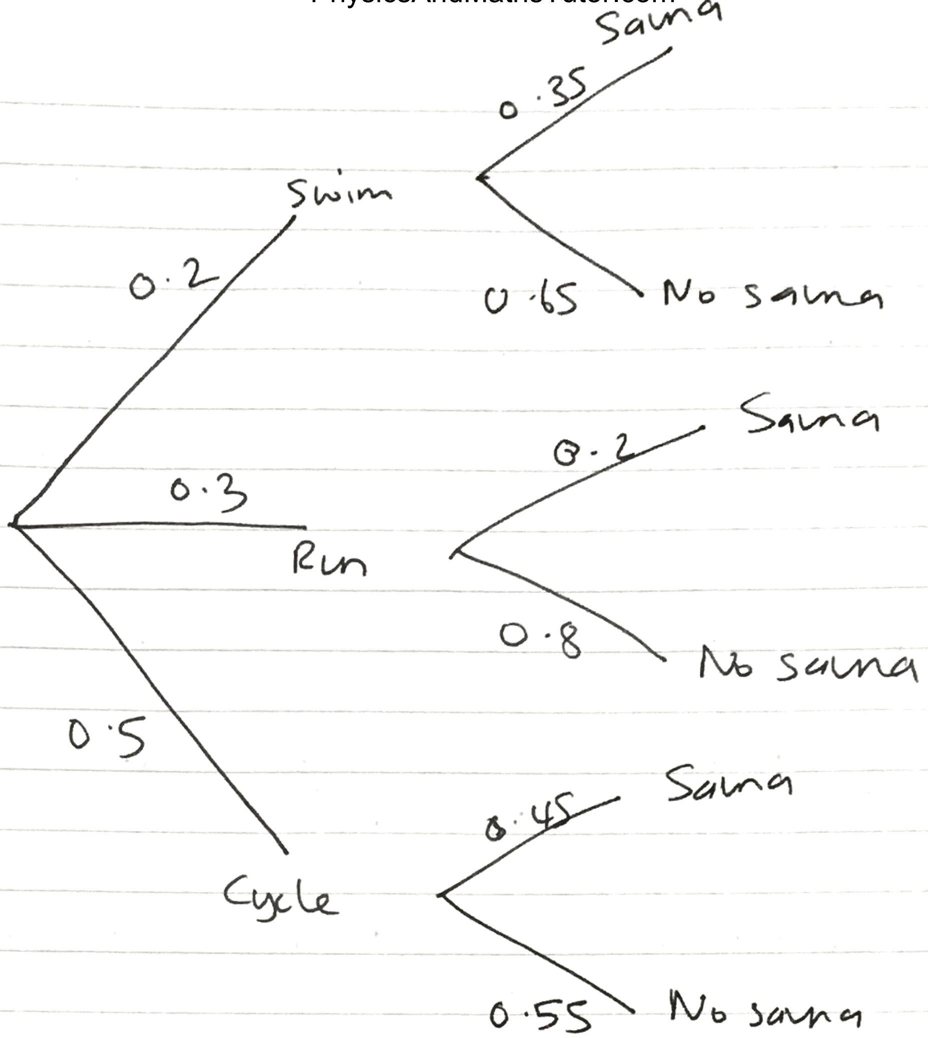
e) $E(A) = 1.624$

$E(A) + 1 = 2.624$

$F[E(A)+1] = P(A \leq 2.624) = P(A=1) + P(A=2)$

$= \boxed{0.84}$

(Q5a)



b) $P(\text{Swim} \cap \text{Sauna}) + P(\text{Run} \cap \text{Sauna}) + P(\text{cycle} \cap \text{sauna})$

$= 0.2 \times 0.35 + 0.3 \times 0.2 + 0.5 \times 0.45$

$= \boxed{0.355}$

c) $P(\text{swimming} \mid \text{used sauna}) = \dfrac{P(\text{swimming and Sauna})}{P(\text{sauna})}$

$= \dfrac{0.2 \times 0.35}{0.355} = \boxed{0.197}$

d) $P(\text{swimming} \mid \text{no sauna}) = \dfrac{P(\text{swimming \& no sauna})}{P(\text{no sauna})}$

from (a)

$= \dfrac{0.2 \times 0.65}{1 - 0.355} = \boxed{0.202}$

Q6a) [REFER TO MARK SCHEME]

b) The scatter diagram from (a) should show the points lying in (approx.) a straight line.

c) $b = \dfrac{S_{xy}}{S_{xx}}$: $S_{xy} = 20615 - \dfrac{(\Sigma x)(\Sigma y)}{8}$

$= 20615 - \dfrac{(260 \times 589)}{8} = 1472.5$

$S_{xx} = 9500 - \dfrac{(260)^2}{8} = 1050$

$\therefore b = \dfrac{1472.5}{1050} = 1.40$ //

$a = \bar{y} - b\bar{x} = \left(\dfrac{589}{8}\right) - (1.40)\left(\dfrac{260}{8}\right) = 28.0$ //

$\Rightarrow \boxed{y = 28.0 + 1.40x}$

d) a = y when x = 0 : a is the temperature generated in the shoulder of the tyre when the vehicle is at rest.

and b is the amount by which the temp. increases when the speed is increased by 1mph.
[ie surrounding temperature increases by 1.40°C when the speed is increased by 1mph.]

e) $x = 50$: $y = 28 + 1.4(50) = 98$ //

There is not a 'perfect' positive correlation. The line calculated is a 'best fit' line so values obtained are only estimates.

f) For 12mph : this would be OK as 12 is just slightly below the lowest x-value in the data (15).

For 85mph : this speed is a lot higher than the highest value in the data (50mph). So using the regression eqn. would result in an extrapolated value which is unreliable.