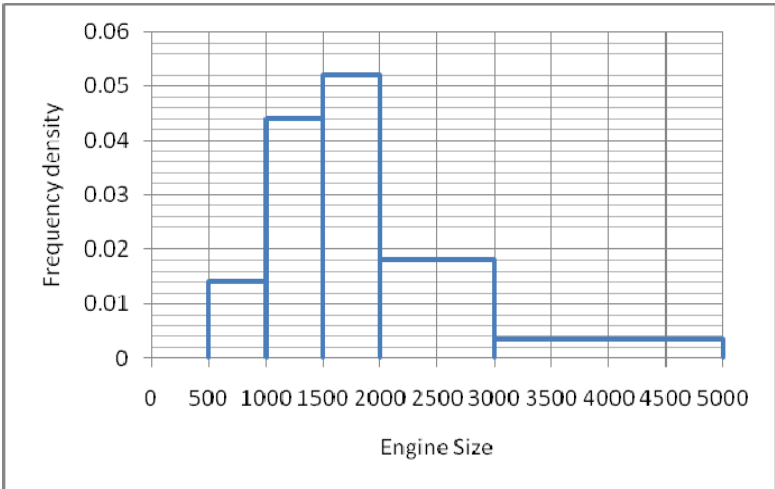


Question		Answer	Marks	Guidance
1	(i)	$4 + \frac{1}{2} \text{ of } 18 = 4 + 9 = 13$	M1 A1 [2]	For $\frac{1}{2}$ of 18 cao 13/100 gets M1A0
	(ii)	(Median) = 50.5 th value $\text{Est} = 140 + \left(\frac{25.5}{29}\right) \times 5 \quad \text{or} = 140 + \left(\frac{50.5 - 25}{54 - 25}\right) \times 5$ $= 144.4$	M1 M1 A1 [3]	For 50.5 seen For attempt to find this value SC2 for use of 50 th value leading to $\text{Est} = 140 + (25 / 29 \times 5) = 144.3$ (SC1 if over-specified) or $\text{Est} = 145 - \left(\frac{3.5}{29}\right) \times 5 = 144.4$ NB no marks for mean = 144.35 NB Watch for over-specification

Question		Answer				Marks	Guidance																									
1	(iii)	<table border="1"> <thead> <tr> <th>Height</th> <th>Fre</th> <th>Group width</th> <th>Frequency density</th> </tr> </thead> <tbody> <tr> <td>$125 \leq x \leq 140$</td> <td>25</td> <td>15</td> <td>1.67</td> </tr> <tr> <td>$140 < x \leq 145$</td> <td>29</td> <td>5</td> <td>5.80</td> </tr> <tr> <td>$145 < x \leq 150$</td> <td>24</td> <td>5</td> <td>4.80</td> </tr> <tr> <td>$150 < x \leq 160$</td> <td>18</td> <td>10</td> <td>1.80</td> </tr> <tr> <td>$160 < x \leq 170$</td> <td>4</td> <td>10</td> <td>0.40</td> </tr> </tbody> </table>				Height	Fre	Group width	Frequency density	$125 \leq x \leq 140$	25	15	1.67	$140 < x \leq 145$	29	5	5.80	$145 < x \leq 150$	24	5	4.80	$150 < x \leq 160$	18	10	1.80	$160 < x \leq 170$	4	10	0.40	M1	For fd's - at least 3 correct	M1 can be also be gained from freq per 10 – 16.7, 58, 48, 18, 4 (at least 3 correct) or freq per 5 – 8.35, 29, 24, 9, 2 for all correct. If fd not explicitly given, M1 A1 can be gained from all heights correct (within one square) on histogram (and M1A0 if at least 3 correct)
		Height	Fre	Group width	Frequency density																											
$125 \leq x \leq 140$	25	15	1.67																													
$140 < x \leq 145$	29	5	5.80																													
$145 < x \leq 150$	24	5	4.80																													
$150 < x \leq 160$	18	10	1.80																													
$160 < x \leq 170$	4	10	0.40																													
A1	Accept any suitable unit for fd such as eg freq per cm. correct to at least one dp allow 1.66 but not 1.6 for first fd																															
						G1	linear scales on both axes and label on vertical axis	Linear scale and label on vertical axis IN RELATION to first M1 mark ie fd or frequency density or if relevant freq/10, etc (NOT eg fd/10). However allow scale given as fd×10, or similar Accept f/w or f/cw (freq/width or freq/class width) Can also be gained from an accurate key																								
						W1	width of bars	G0 if correct label but not fd's. Must be drawn at 125, 140 etc NOT 124.5 or 125.5 etc NO GAPS ALLOWED																								
						H1	height of bars	Must have linear scale. No inequality labels on their own such as $125 \leq S < 140$, etc but allow if a clear horizontal linear scale is also given. Ignore horizontal label.																								
								Height of bars – must be linear vertical scale. FT of heights dep on at least 3 heights correct and all must agree with their																								

Question			Answer	Marks	Guidance													
				[5]	fds If fds not given and at least 3 heights correct then max M1A0G1W1H0 Allow restart with correct heights if given fd wrong (for last three marks only)													
1	(iv)	4 boys 0.6×15 = 9 girls So 5 more girls		M1 A1 A1 [3]	For 0.6×15 For 9 girls cao Or $45 \times 0.2 = 9$ (number of squares and 0.2 per square)													
	(v)	Frequencies and midpoints for girls are <table border="1" style="margin-left: 20px;"> <tr> <td>Height</td> <td>132.5</td> <td>142.5</td> <td>147.5</td> <td>155</td> <td>167.5</td> </tr> <tr> <td>Frequency</td> <td>18</td> <td>23</td> <td>31</td> <td>19</td> <td>9</td> </tr> </table> So mean = $\frac{(132.5 \times 18) + (142.5 \times 23) + (147.5 \times 31) + (155 \times 19) + (167.5 \times 9)}{100}$ $= \frac{(2385) + (3277.5) + (4572.5) + (2945) + (1507.5)}{100}$ = 146.9 (Exact answer 146.875)	Height	132.5	142.5	147.5	155	167.5	Frequency	18	23	31	19	9		B1 B1 M1 M1* Dep on M1 A1	For at least three frequencies correct At least three midpoints correct For attempt at $\sum xf$ For division by 100 Ca NB Watch for over-specification	No further marks if not using midpoints For sight of at least 3 xf pairs Allow answer 146.9 or 147 but not 150 NB Accept answers seen without working (from calculator) Use of 'not quite right' midpoints such as 132.49 or 132.51 etc can get B1B0M1M1A0
Height	132.5	142.5	147.5	155	167.5													
Frequency	18	23	31	19	9													
				[5]														

Question		er				Marks	Guidance
2	(i)	Engine size	Frequency	Group width	Frequency density	M1	At least 4 fds correct for M1 M1 can be also be gained from freq per 1000 – 14, 44, 52, 18, 3.5 (at least 4 correct) and A1 for all correct or freq per 500 - 7, 22, 26, 9, 1.75 Accept any suitable unit for fd, eg freq per 1000, BUT NOT FD per 1000 Allow fds correct to at least three dp If fd not explicitly given, M1 A1 can be gained from all heights correct (within one square) on histogram (and M1A0 if at least 4 correct) Allow restart with correct heights if given fd wrong
		$500 \leq x \leq 1000$	7	500	0.014		
		$1000 < x \leq 1500$	22	500	0.044		
		$1500 < x \leq 2000$	26	500	0.052		
		$2000 < x \leq 3000$	18	1000	0.018		
		$3000 < x \leq 5000$	7	2000	0.0035		
							
		<p>INCORRECT DIAGRAMS: Frequency diagrams can get M0, A0, G0, G1, G0 MAXIMUM Thus frequency density = frequency × width, frequency/midpoint etc gets MAX M0A0G0G1G0 Frequency polygons MAX M1A1G0G0G0</p>				A1 G1(L1)	For fd's all correct linear scales on both axes and label on vertical axis Label required on vert axis IN RELATION to first M1 mark ie fd or frequency density or if relevant freq/1000, etc (NOT fd/1000, but allow fd×1000, etc) Accept f/w or f/cw (freq/width or freq/class width) Ignore horizontal label and allow horizontal scale to start at 500 Can also be gained from an accurate key
						G1(W1)	Width of bars Must be drawn at 500, 1000etc NOT 499.5 or 500.5 etc NO GAPS ALLOWED Must have linear scale. No inequality labels on their own such as $500 \leq S < 1000$, etc but allow if a clear horizontal linear scale is also given.

Question		er	Marks	Guidance
			G1(H1) [5]	Height of bars FT of heights <i>dep</i> on at least 3 heights correct and all must agree with their fds If fds not given and one height is wrong then max M1A0G1G1G0 – visual check on y (within one square) –no need to measure precisely
2	(ii)	Do not know exact highest and lowest values so cannot tell what the midrange is. OR No and a counterexample to show it may not be 2750 OR (500 + 5000) / 2 = 2750. But very unlikely to be absolutely correct but probably close to the true value. Some element of doubt needed. Allow 'Likely to be correct'	E1 [1]	Allow comment such as 'Highest value could be 5000 and lowest could be 500 therefore midrange could be 2750' NO mark if incorrect calculation Sight of 1750 AND 3000 (min and max of midrange) scores E1
	(iii)	Mean = $\frac{(750 \times 7) + (1250 \times 22) + (1750 \times 26) + (2500 \times 18) + (4000 \times 7)}{80}$ $= \frac{151250}{80} = 1891$ $\Sigma x^2 f = (750^2 \times 7) + (1250^2 \times 22) + (1750^2 \times 26) + (2500^2 \times 18) + (4000^2 \times 7)$ $= 3937500 + 34375000 + 79625000 + 112500000 + 112000000$ $= 342437500$ $S_{xx} = 342437500 - \frac{151250^2}{80} = 56480469$ $s = \sqrt{\frac{56480469}{79}} = \sqrt{714943} = 846$ Only an estimate since the data are grouped.	M1 A1 M1 A1 E1 indep [5]	For midpoints (at least 3 correct) No marks for mean or sd unless using midpoints Answer must NOT be left as improper fraction CAO Accept correct answers for mean (1890 or 1891) and sd (850 or 846 or 845.5) from calculator even if eg wrong S_{xx} given For sum of at least 3 correct multiples fx^2 Allow M1 for anything which rounds to 342400000 Only penalise once in part (iii) for over specification, even if mean and standard deviation both over specified. Allow SC1 for RMSD 840.2 or 840 from calculator Or for any mention of midpoints or 'don't have actual data' or 'data are not exact' oe

Question		er	Marks	Guidance
2	(iv)	$\bar{x} - 2s = 1891 - (2 \times 846) = 199$ Allow 200 $\bar{x} + 2s = 1891 + (2 \times 846) = 3583$ Allow 3580 or 3600 So there are probably some outliers	M1 A1 E1 [3]	For either. FT any positive mean and their positive sd/rmsd for M1 Only follow through numerical values, not variables such as s , so if a candidate does not find s but then writes here 'limit is $40.76 + 2 \times \text{standard deviation}$ ', do NOT award M1 No marks in (iv) unless using $\bar{x} + 2s$ or $\bar{x} - 2s$ For both (FT) Do NOT penalise over specification here as it is not the final answer Must include an element of doubt Dep on upper limit in range 3000 – 5000 Allow comments such as 'any value over 3583 is an outlier' Ignore comments about possible outliers at lower end.
	(v)	Number of cars over $2000 \text{ cm}^3 = 25/80 \times 2.5 \text{ million} = 781250$ So duty raised = $781250 \times \text{£}1000 = \text{£}781 \text{ million}$	M1 M1 indep A1 [3]	For $25/80 \times 2.5 \text{ million}$ or $(18+7)/80 \times 2.5 \text{ million}$ For something $\times \text{£}1000$ even if this is the first step CAO NB $\text{£}781250000$ is over specified so only 2/3
	(vi)	Because the numbers of cars sold with engine size greater than 2000 cm^3 might be reduced due to the additional duty.	E1 [1]	Allow any other reasonable suggestion Condone 'sample may not be representative' Allow 'sample is not of NEW cars'

3	(i)	$\text{Percentage} = \frac{40}{200} \times 100 = 20$	M1 A1 [2]	For 40 seen or implied CAO	
	(ii)	Median = 5.2 kg Q1 = 4.2 Q3 = 5.8 Inter-quartile range = $5.8 - 4.2 = 1.6$	B1 B1 B1 [3]	For Q1 or Q3 For IQR	Allow 4.2 to 4.3 for Q1 Dep on both quartiles correct

Question		er	Marks	Guidance	Additional Guidance
3	(iii)	<p>Lower limit $4.2 - (1.5 \times 1.6) = 1.8$ Upper limit $5.8 + (1.5 \times 1.6) = 8.2$ So there are one or more outliers (if any lamb weighs more than 8.2 kg)</p> <p>Should not be disregarded because: ‘Nothing to suggest they are not genuine items of data’ Allow other convincing reasons such as very few so will not make much difference</p>	<p>B1 B1 E1</p> <p>E1</p> <p>[4]</p>	<p>For 1.8 ft For 8.2 ft Dep on their 1.8 and 8.2 Allow any number of outliers ≤ 5</p> <p>Indep Must give reason.</p>	<p>Any use of <u>median</u> ± 1.5 IQR scores B0 B0 E0 E0 if say some outliers at bottom end, unless lower limit > 2.0 If FT leads to limits above 9.0 <u>and</u> below 2.0 then E0 No marks for ± 2 or 3 IQR With 4.3 and 5.8 lower = 2.05 and upper = 8.05 In this part FT their values from (ii) if sensibly obtained but not from location ie 12.5, 37.5 No marks for use of mean $\pm 2s$</p>
	(iv)	<p>Median for Welsh Mountain = 3.6 IQR for Welsh Mountain = 0.8 Welsh Mountain lambs have lower average weight than crossbred</p> <p>Welsh Mountain lambs also have lower variation in weight than crossbred</p>	<p>B1 B1 E1 indep</p> <p>E1 indep</p> <p>[4]</p>	<p>Must imply average or CT, not just median. Allow generally lighter Must imply spread or variation, not just IQR or range Allow correct comment on consistency</p>	<p>FT their medians FT their IQRs Can get max B1B0E1E1 for use of range</p>
	(v)	<p>Median unchanged IQR unchanged OR range or spread increased</p>	<p>E1 E1</p> <p>[2]</p>	<p>even if used IQR in (iv)</p>	<p>E2 for ‘Both comparisons remain the same’ E1 for ‘the range remains smaller’</p>

Question		Answer	Marks	Guidance	Additional Guidance
3	(vi)	$P(\text{Crossbred} > 3.9) = \frac{165}{200}$ $P(\text{Welsh Mountain} > 3.9) = \frac{1}{4}$ $P(\text{Both} > 3.9) = \frac{165}{200} \times \frac{1}{4} = \frac{165}{800} = \frac{33}{160} = 0.206$	B1 B1 M1 A1 [4]	For product of their probabilities, provided one is correct CAO	Allow 162 to 165 out of 200 Allow answers in range 0.2025 to 0.20625 with correct working