adamjee maths

A-LEVEL MATHS / EDEXCEL / FREE RESOURCES

Data presentation and interpretation

Mark scheme

13 questions

Exam questions and marking instructions remain credited to their board.

Contents

1 June 2025 - Paper 3 Statistics - Q2

2 June 2025 - Paper 3 Statistics - Q5

3 June 2024 - Paper 3 Statistics - Q3

4 June 2024 - Paper 3 Statistics - Q5

5 June 2023 - Paper 3 Statistics - Q3

6 June 2023 - Paper 3 Statistics - Q4

7 June 2023 - Paper 3 Statistics - Q6

8 October 2021 - Paper 3 Statistics - Q5

9 October 2020 - Paper 3 Statistics - Q3

10 October 2020 - Paper 3 Statistics - Q5

11 June 2019 - Paper 3 Statistics - Q2

12 June 2019 - Paper 3 Statistics - Q3

13 June 2018 - Paper 3 - Q4

2(a)	$\left[\bar{x} = \frac{6612}{120}\right] =$ **<u>55.1</u>** (seconds)	B1	1.1b
		(1)	
2(b)	$[\sigma_x =]\sqrt{\frac{364\,902}{120} - (\text{"}55.1\text{"})^2}$ <u>or</u> $\sqrt{4.84}$	M1	1.1b
	= **<u>2.2</u>** (seconds)	A1	1.1b
		(2)	
2(c)	Coach *B* because standard deviation is greater	M1	2.4
	This means *B* has a greater range **and** therefore more extreme values	A1ft	2.2b
		(2)	
			(5 marks)

Notes:
Ignore units in this question

(a)	B1	for 55.1 or $55\frac{1}{10}$ only (may be seen in the question) ISW e.g. 55.1 followed by 55 etc Do not accept $\frac{551}{10}$ that scores B0
(b)	M1	for a correct expression (values seen not just a formula)(including $\sqrt{}$)
	A1	for 2.2 <u>or</u> $\frac{11}{5}$ o.e. (allow *s* = 2.209… so awrt 2.21) Correct answer is 2 marks
(c)	**Ignore comments about different number of runners. Don't need to state means are the same.** **If there is no answer given for (b) then no marks can be scored in (c)**	
"2.2"<3.6	B1	for coach *B* **and** states s.d. of *B* is greater o.e. Must compare. Allow implied choose *B* e.g. *B* has a greater range <u>or</u> *B*s values are more spread out <u>or</u> 3.6 > "2.2" <u>or</u> $\sigma_B >$ "2.2" May say s.d. of *A* is smaller which is equivalent to saying s.d. of *B* is larger
	dB1	(dep on 1st B1 scored) for clear reason why greater s.d. means choose coach *B* Idea of greater range **and** more extreme values (faster times) in *B* e.g. *B* has more (<u>or</u> greater chance of <u>or</u> greater % of) runners at lower times (may compare mean – 1 or 2 or 3 standard deviations to show that *B* has the quicker time) e.g. coach *A* (52.9) > (51.5) coach *B* so coach *B* likely to be quickest
	NB	**"**coach *B* since *B* has greater range /s.d. and so more extreme values" scores B1B1
SC **"2.2">3.6**	**(A0 in (b)** and answer > 3.6)	
	B1	equivalent reasons as above but choose *A*
	dB1	same rules as above but applied to *A*

Question	Scheme	Marks	AOs

		(2)	
5(b)	[Let F be the height of a women from the club] P(F < 170) = 0.40 $\Rightarrow \frac{170-\mu}{\sigma} = -0.2533$ [calc gives – 0.253347…]	M1	3.1b
	P(F > 175) = 0.15 $\Rightarrow \frac{175-\mu}{\sigma} = 1.0364$ [calc gives 1.036433…]	M1	3.4
	Solving: $5 = \sigma(1.0364 + 0.2533)$	M1	1.1b
	σ = 3.87687… awrt **3.88**	A1	1.1b
	μ = 170.9820… awrt **171**	A1	3.2b
		(5)	
5(c)	[P(170 < F < 175) =] 1 – (0.40 + 0.15) = 0.45 (allow awrt 0.45) [P(170 < M < 175) =] 0.047282… = (allow awrt 0.047)	M1	3.4
	[P(170 < F < 175) × P(170 < M < 175) =] "0.45"×"0.047" or 0.021…	A1ft	1.1b
	So P(both) = 0.021277… awrt **0.021**	A1	1.1b
		(3)	
			(10 marks)

Notes:

(a)(i)	B1	for awrt 0.270 (accept just 0.27)
(ii)	B1	for awrt 0.607
(b)	M1	for standardising 170 with μ and σ and set equal to a z value $0 < \|z\| < 0.3$
	M1	for standardising 175 with μ and σ and set equal to a z value $1 < \|z\| < 2$
	M1	for solving their 2 linear equations, reducing to a linear equation in one variable This can be implied by two answers that are correct to 2sf.
	To award the A marks they must have scored 1st M1 and 2nd M1	
	A1	for = σ awrt 3.88
	A1	for = μ awrt 171 (Must not be from incorrect working e.g. $\sigma < 0$)
	Ans only	We need to see the linear equations in μ and σ to start awarding marks. If 1st M1 and 2nd M1 are scored and correct answers are seen we can award 5 marks.
	MR	**(15% as 0.015)** 2nd M1 for z = 2.17…, 3rd M1 implied by σ = 2.06… and μ = 170.5 … i.e. Maximum M1M1M1A0A0
(c)	M1	for correct probability for women or correct probability for men or clear intention to multiply P(170 < F < 175) by P(170 < M < 175)
	A1ft	for using P(170 < F < 175) × P(170 < M < 175) with one probability correct to 2sf
	A1	for awrt 0.021. Correct answer only is 3 marks.
	MR	**(if used in (b))** P(170< F < 175) = 0.585 (or awrt 0.58 or 0.59) answer awrt 0.027 or 0.028 All 3 marks can be scored here.

Qu 3	Scheme	Marks	AO
(a)	Rain[fall] (allow [Mean] Windspeed)	B1 **(1)**	1.2
(b)	[$\bar{x}$ =] 15.2239… = awrt **15.2**	B1	1.1b
	$\sigma_x = \sqrt{\frac{44\,695.4}{184} - "15.22.."^2}$ or $\sqrt{11.1(422...)}$	M1	1.1b
	= 3.33800… awrt **3.34**	A1 **(3)**	1.1b
(c)	Mean is higher than average **OR** a summer/spring month If they say winter/autumn they must explain that these are hotter months for Perth. [Perth is southern hemisphere or Australia so latest available] month is **Oct**	M1 A1cso **(2)**	2.4 2.2b
		(6 marks)	

	Notes
	Answers may appear next to the question.
(a)	B1 for Rain[fall] or precipitation (e.g. allow Daily Total [or Mean or max] Rainfall etc) (or allow Mean Windspeed or just "windspeed" BUT not max windspeed or "gust") If they give more than one answer we take the **last one**. (NB Actual windspeed mean is 8.2, sd 2.38. No other quantitative variables available)
(b)	B1 for awrt 15.2 (Do not accept fractions or mixed numbers) M1 for a correct expression including square root (ft their mean) May be implied by an answer of 3.3 or better. A1 for awrt 3.34 [Allow s = 3.3471… i.e. awrt 3.35 if correct formula/expression is seen]
	If answer in (b)(i) > 19.4 and an attempt is made in (c) please send to review.
(c)	M1 for a reason mentioning that mean or temperature is higher (o.e.) e.g. it is a warmer/hotter month is OK or sight of 19.4 > (their) 15.2 Only ft their 15.2 if it is less than 19.4 **OR** suggesting a summer/spring month. Ignore incorrect statements that are irrelevant or don't contradict For incorrect statements that contradict score M0 A1cso dep on M1 scored for inferring October Must choose just October not a range like August~October (NB actual mean for Sep is 15.6 and sd 3.19 and this scores A0) Can accept for example "high mean so December" for M1A0
SC	M1A1 for "October since Perth is in the southern hemisphere/Australia" M1A0 for "Sep or Nov or Dec or Jan or Feb **and** "Perth is in the southern hemisphere/Australia" M0A0 just "Perth is in the southern hemisphere/Australia" without a month M0A0

Qu 5	Scheme	Marks	AO
(a)	[P($H > 1.6$) =] 0.091211… = awrt **0.0912**	B1 **(1)**	1.1b
(b)	Need H and T to be independent or events $\{H > 1.6\}$ and $\{T < 300\}$ are independent	B1 **(1)**	2.4
(c)	[P($T < 300$) =] 0.124(2816…) Prob both is: "0.0912..."×"0.124..." = 0.011335… = awrt **0.0113**	M1 M1 A1 **(3)**	3.4 1.1b 1.1b
(d)	$\frac{16.3-\mu}{\sigma} = -0.5244(0051...)$, $\frac{29-\mu}{\sigma} = 1.2816$ (calc: 1.28155156…) e.g. $29-16.3 = \sigma$("1.2816"−"−0.5244") $\sigma = 7.032115...$ = awrt **7.03** $\mu = 19.9876...$ = **19.95** „ **μ** „ **20.0**	M1M1 M1 A1 A1 **(5)**	3.1a 1.1b 1.1b 1.1b 3.2a
		(10 marks)	

	Notes
(a)	B1 for awrt 0.0912 (from calculator)
(b)	B1 for a suitable reason mentioning or implying H and T are independent Allow: e.g. "they"/ "each event"/ "P(H) and P(T)"/ "the variables" **and** "independent" B0 for "the results" /"the values" are independent. Ignore other comments that are not incorrect or contradictory.
(c)	1st M1 for using model for T to attempt to find P($T < 300$) e.g. sight of 0.124 or better or sight of $\pm\left(\frac{300-330}{26}\right)$ or $\pm\left(\frac{5-5.5}{0.433...}\right)$ or $Z = \pm 1.15(3…)$ 2nd M1 for multiplying their two probabilities together ft part (a) and their P($T < 300$) provided both values are probabilities. NB M0M1 is possible here A1 for awrt 0.0113 [Correct answer with no incorrect working 3/3]
(d)	1st M1 for standardising 16.3 and setting equal to z value where $0.5 < \|z\| < 0.6$ 2nd M1 for standardising 29 and setting equal to z value where $1 < \|z\| < 1.5$ 3rd M1 dep on 1st or 2nd M1 for solving their two linear eq'ns – reach an eq'n in one variable May be implied by sight of $\sigma = 7$ (or better) or $\mu = 20$ (or better) **For 1st A mark we must also see one of – 0.5244 or 1.2816 (or better) used in their equ'ns OR both z values correct to 3dp i.e. – 0.524 and 1.282** 1st A1 for σ = awrt 7.03 (but see 3rd case below) 2nd A1 for μ = in [19.95, 20.0] (i.e shouldn't see something rounding down to 20.0) Allow 20 from equations with suitable z values (see examples below)
NB	Use of – 0.524 and 1.28 [would give 7.0399… and 19.988…] and scores M3A0A1 Use of – 0.524 and 1.2816 [would give 7.033… and 19.99…] and scores M3A1A1 Use of – 0.5244 and 1.28 [would give 7.038… and 19.99 …] and scores M3**A1**A1 Both z values correct to 3dp i.e. – 0.524 and 1.282 [should give 7.032 and 19.984] scores A1A1

Qu 3	Scheme	Marks	AO
(a)	Need to replace tr with a numerical value Value of tr is between 0 and 0.05 suggest using e.g 0.025 , 0 <u>or</u> value „ 0.05	M1 A1 **(2)**	1.2 1.1b
(b)(i)	$\left[\bar{x} = \frac{389.3 \sim 390.8}{184}\right] = 2.119\ldots$ awrt **<u>2.12</u>** allow $\frac{195}{92}$ or $2\frac{11}{92}$	B1	1.1b
(ii)	$\left[\sigma =\right]\sqrt{\frac{(\text{awrt})4336}{184} - "\bar{x}^2"}$ <u>or</u> allow $\left[\sigma^2 =\right]\frac{(\text{awrt})4336}{184} - "\bar{x}^2"$ <u>or</u> awrt 19.1 $= 4.367\ldots$ awrt **<u>4.37</u>**	M1 A1 **(3)**	1.1b 1.1b
(c)(i)	Only covers May~Oct (so not a suitable sample)	B1	1.1b
(ii)	e.g. Winter months are <u>missing</u> when we'd expect <u>more rain</u> so expect estimate in (b)(i) to be an <u>underestimate</u> (oe)	B1 **(2)**	2.4
		(7 marks)	

Notes

(a) M1 for recognising that tr must be replaced (oe) with a numerical value
The following examples would score M0: The tr values are worth 0 so ignore (not replacing) <u>or</u> must remove outliers <u>or</u> fill gaps in table <u>or</u> make widths the same <u>or</u> need to find mid-points
A1 for using a suitable value: e.g. 0.025 (or allow 0) i.e. any value in [0, 0.05]
(these give $\sum x = 390$ (3sf), use of 0.05 gives 390.8, use of 0 gives 389.3 allow in (b)(i))

(b)(i) B1 for awrt 2.12 <u>or</u> allow simplified fraction or mixed number. B0 for $\frac{390}{184}$

(ii) M1 for a correct expression for standard deviation **or** variance. Allow $\sum x^2$ = awrt 4336
Ignore their label σ or σ^2 Can ft their mean

A1 for awrt 4.37 [Use of *s* gives 4.3791… so for correct use seen allow awrt 4.38]

SC **Using *n* = 155** Allow M1 for expression $\left[\sigma =\right]\sqrt{\frac{(\text{awrt})4336}{155} - "\bar{x}^2"} = \sqrt{21.64...}$ or 4.65…

Part (c) can effectively be marked together.

(c)(i) B1 for a comment mentioning that data is just from May~Oct (so not representative of the whole year).
Just saying "only 184 days so not representative" is B0, **must mention May ~ Oct**

(ii) B1 for comment that <u>missing</u>/<u>winter</u> months expected to have more rain (oe) **and** "underestimate"(oe)
We are looking for all 3 of these ideas here:
1. A statement or implication that missing data is from winter or different months.
2. A suggestion about the rainfall in these months (probably more rain).
3. A statement about the impact on the estimate in (b)(i) <u>equivalent</u> to saying it would be an underestimate or the (actual) mean will be higher.

SC If you see "Leeming or N or NE has <u>less</u> rain in winter months" – please send to review

Qu 4	Scheme	Marks	AO
(a)	[Let N = height from region A; P($N > 180$) =] 0.24937… awrt **0.249**	B1 **(1)**	1.1b
(b)	$H_0 : \mu = 175.4$ $H_1 : \mu \neq 175.4$	B1	2.5
	[S = height from region B] $\bar{S} \sim N\left(175.4, \frac{6.8^2}{52}\right)$ Allow $\sigma^2 =$ awrt 0.889	M1	3.3
	[P($\bar{S} > 177.2$)] = 0.02814…	A1	3.4
	[0.028… > 0.025, Not sig, do not reject H_0]	A1	2.2b
	Insufficient evidence to support student's claim	**(4)**	
(c)	[p-value = $2 \times 0.02814...$ =] 0.05628… in range **0.056~0.06** or **5.6(%)~6(%)**	B1ft **(1)**	1.2
		(6 marks)	

	Notes
(a)	B1 for awrt 0.249
(b)	B1 for both hypotheses correct in terms of μ (**See below for one-tail test**)
	M1 for selecting the correct model, may be implied by standardisation using correct values or may be implied by a correct value in 1st A1 e.g.(Prob =) 0.028 or awrt 0.972, (Z =) 1.9(08..) (CV=) 177.25 Condone use of S (or any other letter) instead of $\bar{S}$ Condone use of $\bar{S} \sim N\left(177.2, \frac{6.8^2}{52}\right)$ but **this will lose 2nd A mark**
	1st A1 for probability of awrt 0.028 (allow 0.03 if P($\bar{S} > 177.2$) is seen) Condone 1 – 0.02814 … = 0.9718…(awrt 0..972) **only if clearly compared with 0.975**
ALT	Allow Z = 1.9(088…) and comparison with 1.96 (or better: calc gives 1.95996…) or CR of $\lceil \bar{S} \rceil$… 177.248…(awrt 177.25) Allow $\lceil \bar{S} \rceil >$ 177.248…(awrt 177.25) Implied by diagram or correct interpretation of inequality with their CV (Ignore any attempt at a lower CR for $\bar{S}$)
	2nd A1 (dep on 1st A1 and use of correct model. Use of N(177.2, …) scores A0) for a conclusion using context: e.g. does not support student's claim or e.g. insufficient evidence of a difference in heights Do not allow 2nd A mark for contradictory statements e.g. "significant" so "no support for claim"
(c)	B1ft for answer in range 0.056~0.06 or 5.6%~6% (Ranges are inclusive, condone missing %) (can ft their probability, provided < 0.5, from part (b) but not 0.025 leading to 5%)
NB	**One-tail test [Max of 3/5 for (b) and (c)]** In (b) B0 (hypotheses) M1(model as above) 1st A1[for probability or Z compared with 1.6449 or CR $\lceil \bar{S} \rceil$… or > 176.95… (awrt 177)] 2nd A1 for conclusion in context that supports claim **or** "heights of men from B is different from/greater than from A" In (c) B0

Qu 6	Scheme	Marks	AO
(a)	$2\times4.2\,,\,4\times4\,,\,4\times3.5\,,10\times1$ $(=8.4+16+14+10=48.4)$	M1	1.1b
	[So P$(10 < T < 30)$ =] $\left[\dfrac{48.4}{90}\right]=\dfrac{121}{225}=0.53777...$ **<u>0.53~0.54</u>** (2sf OK)	A1	1.1b
		(2)	
(b)	(Not suitable as) data is not symmetric <u>or</u> is skew (normal is symmetric) ("Even" distribution or a diagram <u>on its own</u> is not enough so B0)	B1 **(1)**	2.4
(c)	$\int xe^{-x}\ (dx)=\int x\,d(-e^{-x})$	M1	2.1
	$=\left[-xe^{-x}\right]-\int\left(-e^{-x}\right)(dx)$ $(+c)$	A1	1.1b
	$\int_0^n xe^{-x}\ (dx)=\left[-xe^{-x}-e^{-x}\right]_0^n=\left(-ne^{-n}-e^{-n}\right)-\left[-(0)-1\right]$	dM1	1.1b
	$=\underline{1-(n+1)e^{-n}}$ (*)	A1cso*	1.1b
		(4)	
(d)	Require area = 90 i.e. $k\int_{(0)}^{(n)} xe^{-x}\,dx=90$ (ignore limits)	M1	3.1a
	Using the result in part (c) with $n=4$ gives $k\left[1-5e^{-4}\right]=90$	M1	2.1
	(k =) **<u>99</u>**(.0729…) (*****)	A1cso*	1.1b
		(3)	
(e)(i)	[P$(10 < T < 30)$ =] 0.64863… awrt **<u>0.649</u>**	B1	1.1b
		(1)	
(ii)	[No. of patients =] $(99)\left[\left(1-4e^{-3}\right)-\left(1-2e^{-1}\right)\right]$ $(=53.1..)$	M1	3.4
	Prob $=\dfrac{0.5366...\times99}{90}=0.59027...$ [or 0.5907...] = awrt **<u>0.590</u>** or **<u>0.591</u>**	A1	3.2a
		(2)	
(f)	eg Patients might stay longer than 40 hours (Can ignore other comments unless clearly contradictory.)	B1 **(1)**	3.5b
		(14 marks)	

	Notes
(a)	M1 for an attempt to find the number between 10 and 30 (2 correct products or 48 or 48.4 seen) A1 for 2sf answer in [0.53 ~ 0.54] NB use of 48 gives 0.5333… [Correct ans implies 2/2]
(b)	B1 for a comment suggesting not suitable based on (lack of) symmetry <u>or</u> "not bell shaped"
(c)	1st M1 for attempting integration by parts in right direction. Must have $u = x$ and $v = \pm e^{-x}$ 1st A1 for a correct first step, correct first integration and expression for second integral 2nd dM1 (dep on 1st M1) for all integration attempted and some use of at least one limit * 2nd A1 for cso with no incorrect working seen. Minimum is correct int and use of limits seen.
(d)	1st M1 for realising need area under the curve (implied by the integral) = 90 2nd M1 for use of (c) with $n = 4$ and set = 90 May be implied by sight of 99.07… or better * A1cso for $k = 99$ or awrt 99.1 **NB** Allow use of $k = 99$ and show area = awrt 89.9 with a conclusion to score 3/3
(e)(i)	B1 for awrt 0.649
(ii)	M1 for use of (c) with $n = 1$ and $n = 3$ Don't need the 99. Implied by sight of awrt 0.54 A1 for awrt 0.590 or awrt 0.591 Allow 0.59 from correct working seen.
(f)	B1 eg for comment, in context, about the upper limit for <u>time</u> (<u>t</u> or <u>x</u>)(time/hour may be implied)

Notes on Question 5

The question essentially uses the definition of P($A \mid B$) given in the formula booklet.

In particular $\mathrm{P}\left(S \mid \{X = x\}\right) = \dfrac{\mathrm{P}\left(S \cap \{X = x\}\right)}{\mathrm{P}\left(X = x\right)}$ [1]

The first "blob" tells us that $\mathrm{P}\left(S \mid \{X = x\}\right) = \dfrac{k}{x}$ where k is a constant.

The second "blob" tells us that $\mathrm{P}\left(S \cap \{X = x\}\right)$ is the same for all x so $\mathrm{P}\left(S \cap \{X = x\}\right) = V$ where V is a constant.

Using these results in [1] gives $\dfrac{k}{x} = \dfrac{V}{\mathrm{P}(X = x)}$ [2]

Line 1 of MS for part (a) uses $V = \mathrm{P}\left(X = x\right) \times \dfrac{k}{x}$ for $x = 50$ and $x = 80$

Line 2 of MS for part (a) uses [2] with $x = 50$ and $x = 80$

Other implications

Equation [1] can be rearranged to give $\mathrm{P}\left(X = x\right) = x \times \dfrac{V}{k}$ [3]

So when $a + b + c + d = 1$ is used this gives $1 = \dfrac{V}{k}\left(20 + 50 + 80 + 100\right)$ or $\dfrac{V}{k} = \dfrac{1}{250}$ [4]

In particular if we use this relationship in [3] the probabilities a, b, c and d can simply be written down for example $b = \dfrac{50}{250}$ as given in the **NB** in the notes on the MS.

The point is that k and V will vary according to equation [4] but as part (c) shows there are some restrictions on the values k, and therefore V, can take.

Since $\dfrac{k}{x}$ is a probability then, ignoring the trivial cases*, $0 < \dfrac{k}{x} < 1$ and the "restricting" value of x is clearly $x = 20$ so $0 < k < 20$ and from [4] we get $0 < V < \dfrac{20}{250} = \dfrac{2}{25} = a$

So the restrictions on k and on V are given by the shortest distance and its associated probability.

* $k = 0$ would say Tisam can never get the ball in the cup no matter what the distance.

$k = 20$ says she always gets the ball in the cup for any distance.

Greg Attwood June 2023

Pearson Education Limited. Registered company number 872828
with its registered office at 80 Strand, London, WC2R 0RL, United Kingdom

Qu 5	Scheme	Marks	AO
(a)	$\left[\text{Let } F \sim \text{N}\left(166.5, 6.1^2\right)\right]$ $\text{P}(F < k) = 0.01 \Rightarrow \dfrac{k-166.5}{6.1} = -2.3263$	M1	3.4
	$k = 152.309\ldots$ **<u>152</u>** or awrt **<u>152.3</u>**	A1 **(2)**	1.1b
(b)	[P($150 < F < 175$) =] 0.914840… awrt **<u>0.915</u>**	B1 **(1)**	1.1b
(c)	P($F > 160 \mid 150 < F < 175$)	M1	3.1b
	$= \dfrac{\text{P}(160 < F < 175)}{\text{P}(150 < F < 175)}$ or $\dfrac{\text{P}(160 < F < 175)}{\text{"(b)"}}$	M1	1.1b
	$= \dfrac{0.7749487...}{\text{"0.91484..."}}$	A1ft	1.1b
	= 0.84708… awrt **<u>0.847</u>**	A1 **(4)**	1.1b
(d)	$\text{H}_0 : \mu = 166.5$ $\text{H}_1 : \mu < 166.5$	B1	2.5
	[Let X = height of female from 2nd country] $\bar{X} \sim \text{N}\left(166.5, \left(\dfrac{7.4}{\sqrt{50}}\right)^2\right)$	M1	3.3
	$\text{P}\left(\bar{X} < 164.6\right) = 0.03472...$	A1	3.4
	[0.0347… < 0.05 so significant <u>or</u> reject H_0] There is evidence to support Mia's belief	dA1 **(4)**	2.2b
		(11 marks)	

	Notes
(a)	M1 for standardising (allow <u>+</u>) with k, 166.5 and 6.1 and set equal to a z value $2.3 < \lvert z \rvert < 2.4$ A1 for 152 or awrt 152.3 **Ans only** 2/2 [Condone poor use of notation e.g. $\text{P}(\frac{k-166.5}{6.1}) = -2.3263$] **Allow percentages instead of probabilities throughout.**
(b)	B1 for awrt 0.915
(c)	1st M1 for interpreting demand as an appropriate conditional probability ($\Rightarrow$ by 2nd M1) 2nd M1 for correct ratio of expressions (can ft their (b) on denominator) ($\Rightarrow$ by 1st A1ft) 1st A1ft for a correct ratio of probs (can ft their "0.9148…" to 3sf from (b) if > 0.775) 2nd A1 for awrt 0.847
(d)	B1 for both correct hypotheses in terms of μ 1st M1 for selecting the correct model (needn't use $\bar{X}$ $\Rightarrow$ by standardisation or 1st A1) 1st A1 for correct use of the correct model i.e. awrt 0.035 (allow 0.04 if $\text{P}\left(\text{"}\bar{X}\text{"} < 164.6\right)$ seen) Condone $\text{P}(\text{"}\bar{X}\text{"} > 164.6) = 0.9652$ or awrt 0.97 <u>only if</u> comparison with 0.95 is made
ALT	**Use of *z* value:** Need to see $Z = -1.8(15\ldots)$ **and** cv of <u>+</u> 1.6449 (allow 1.64 or better) for 1st A1
ALT	**Use of CR or CV for $\bar{X}$:** Need to see "$\bar{X}$"< 164.7786… or CV = … (awrt 164.8) for 1st A1 Condone truncation i.e 164.7 or better 2nd dA1 (**dep on M1A1** only) for a correct inference in context. Must mention <u>Mia's belief</u> **or** <u>mean height of females/women</u> Do NOT award if contradictory statements about hypotheses made e.g. "not sig"
SC	**M0 for** $\bar{X}$ ~N(**164.6**, …) If they achieve p = awrt 0.035 (o.e. with z-value or CV of 166.3) **and** a correct conclusion in context is given score M0A0A1 [and SC for awrt 0.97 > 0.95 case]

Qu 3	Scheme	Marks	AO
(a)	[68 – 7 =] **61** (only)	B1 **(1)**	1.1b
(b)	[25 – 14] = **11**	B1 **(1)**	1.1b
(c)	$\left[\mu \text{ or } \bar{x} = \frac{607.5}{27} = \right]$ = **22.5**	B1 **(1)**	1.1b
(d)	$\sigma = \sqrt{\frac{17\,623.25}{27} - "22.5"^2}$ or $\sqrt{146.4629...}$	M1	1.1b
	= 12.10218… awrt **12.1**	A1 **(2)**	1.1b
(e)	$\mu + 3\sigma = "22.5" + 3 \times "12.1..."$ = awrt 59 so only **one** outlier	B1ft **(1)**	1.1b
(f)	Median increases implies that both values must be > 20	M1	3.1b
	Mean is the same means that $a + b = 45$	M1	1.1b
	So possible values are: e.g. $b = 21$ and $a = 24$ (o.e.)	A1 **(3)**	2.2b
(g)	Both values will be less than 1 standard deviation from the mean and so the standard deviation of all 29 values will be smaller	B1 **(1)**	2.4
		(10 marks)	

Notes

(a) B1 for correctly interpreting the box plot to find the range (more than 1 answer is B0)

(b) B1 for correct understanding of IQR and answer of 11

(c) B1 for 22.5 only (or exact equivalent such as $\frac{45}{2}$). Allow 22 mins and 30 secs.

(d) M1 for a correct expression including square root. Allow $\sqrt{146}$ or better. Ft their mean
A1 for awrt 12.1 NB Allow use of s = 12.3327… or awrt 12.3

(e) B1ft for a correct calculation or value based on their μ and σ and compatible conclusion

(f) 1st M1 Correct start to the problem and a correct statement about the values based on median
Allow if their final two values are both >20
2nd M1 for a correct explanation leading to equation $a + b = 45$ (o.e. e.g. equidistant from mean)
Allow if their final two values sum to 45
A1 for a correct pair of values (both > 20 with a sum of 45) **and** at least some attempt to explain how their values satisfy at least one of the conditions (both > 20 or $a + b = 45$).
Ignore $a =$ or $b =$ labels
NB The values for a and b do not need to be integers.

(g) B1 for a correct explanation.
Must mention that both values are less than 1 sd (ft their answer to (d)) from the mean

Qu 5	Scheme	Marks	AO
(a)	{Let X = time spent, P(X > 15) = } 0.105649… awrt **0.106**	B1 **(1)**	1.1b
(b)	$H_0 : \mu = 10$ $H_1 : \mu > 10$	B1	2.5
	$\bar{X} \sim N\left(10, \left(\frac{4}{\sqrt{20}}\right)^2\right)$; $P(\bar{X} > 11.5)$ = 0.046766… [Condone 0.9532…]	M1;A1	3.3;3.4
	[This is significant (< 5%) so] there is evidence to support the complaint	A1 **(4)**	2.2b
(c)(i)	[P(T < 2) =] 0.1956… awrt **0.196**	B1 **(1)**	1.1b
(ii)	Require $\frac{P(0 < T < 2)}{P(T > 0)} = \frac{0.119119...}{0.923436...}$; = 0.1289955… awrt **0.129**	M1 A1;A1 **(3)**	3.4 1.1bx2
(iii)	The current model suggests **non-negligible** probability of T values < 0 which is impossible	B1 **(1)**	3.5b
(d)	Require t such that P($T > t \mid T > 2$) = 0.5 or P($T < t \mid T > 2$) = 0.5	M1	3.1b
	e.g. $\frac{P(T > t)}{P(T > 2)} = 0.5$; so P($T > t$) = 0.5×[1 – (c)(i)] or P($T > t$) = 0.5×0.8043..	M1; A1ft	1.1b 3.4
	[i.e. P($T > t$) = 0.40… implies] $\frac{t-5}{3.5} = 0.2533$ or P($T < t$) = "0.5978.."	M1	1.1b
	t = 5.886… or from calculator 5.867… so awrt **5.9**	A1 **(5)**	1.1b
		(15 marks)	

	Notes
(a)	B1 for awrt 0.106 (from calculator) [Allow 10.6%]
(b)	B1 for both hypotheses correct in terms of μ. M1 for selection of a correct model (sight or use of correct normal- may not have label $\bar{X}$) 1st A1 for use of this model to get probability allow 0.046~0.047 [Condone awrt 0.953]
ALT	**OR** test statistic z = 1.677… (awrt 1.68) and cv of 1.64 (or better) **or** CR $\bar{X}$ > 11.47.. 2nd A1 (dep on 1st A1 or at least $P(\bar{X} > 11.5) < 0.05$ (o.e.))
SC	for a correct conclusion in context -must mention **complaint**/claim or **time**/mins is > 10 **(M0 for $\bar{X}$ ~N(11.5, …)** for correct probability **and** conclusion (score M0A0A1 on epen)
(c)(i)	B1 for awrt 0.196 (from calculator) [Allow 19.6%]
(ii)	M1 for a correct probability ratio expression (may be implied by 1st A1 scored) 1st A1 for a correct ratio of probabilities (both correct or truncated to 2 dp) 2nd A1 for awrt 0.129
(iii)	B1 for a suitable explanation of why model is not suitable based on negative T values Must say that a **significant** proportion of values < 0 (o.e.) e.g. P(T > 0) should be **closer** to 1 or Difference between P(T < 2| T > 0) and P(T < 2) is **too big** (o.e.)
(d)	1st M1 for a correct conditional probability statement to start the problem or 0.5×P(T > 2) 2nd M1 for correct ratio of probability expressions [Must have P($T > t$) or P(2 < T < t)] 1st A1ft for a correct equation for P($T > t$) (o.e.) ft their answer to part (c)[May be in a diagram] 3rd M1 for attempt to find t (standardising and sight of 0.2533) or prepare to use calc (ft) Arriving at P(T < median) = 1 – 0.5×"their 0.8043" will score 1st 4 marks 2nd A1 for awrt 5.9 Sight of awrt 5.9 and at least one M mark scores 5/5 [Answer only send to review]

Pearson Education Limited. Registered company number 872828
with its registered office at 80 Strand, London, WC2R 0RL, United Kingdom

Question	Scheme	Marks	AOs
2(a)	IQR = 26.6 – 19.4 [= 7.2]	B1	2.1
	19.4 – 1.5 × '7.2' [= 8.6] or 26.6 + 1.5 × '7.2' [= 37.4]	M1	1.1b
	Plotting one upper whisker to 32.5 **and** one lower whisker to 8.6 or 9.1	A1	1.1b
	Plotting 7.6 and 8.1 as the only two outliers	A1	1.1b
		(4)	
(b)	October (since it is the month with the coldest temperatures between May and October in Beijing)	B1	2.4
		(1)	
(c)	$[\sigma =]\sqrt{\frac{4952.906}{184}}$ or e.g. $[\sigma =]\sqrt{\frac{S_{xx}}{n}} = 5.188...$ [=5.19*]	B1cso*	1.1b
		(1)	
(d)	$z = (\pm)\ 1.28(16)$ $[P_{90} =]29.251...$ or $[P_{10} =]15.948...$	B1	3.1b
	2 × 1.2816 × 5.19 '29.251…' – '15.948…'	M1	1.1b
	= awrt **13.3**	A1	1.1b
		(3)	
(e)	Daily mean wind speed/Beaufort conversion since it is qualitative Rainfall since it is not symmetric/lots of days with 0 rainfall	B1 B1	2.4 2.4
		(2)	
			(11 marks)

Notes

(a)

B1: for a correct calculation for the IQR (implied by 10.8 **or** 8.6 **or** 37.4 seen)
M1: for a complete method for either lower outlier limit or upper outlier limit (allow ft on their IQR) (may be implied by the 1st A1 or a lower whisker at 8.6)
A1: both whiskers plotted correctly (allow ½ square tolerance)
A1: only two outliers plotted, 7.6 and 8.1 (must be disconnected from whisker)
NOTE: A fully correct box plot with no incorrect working scores 4/4

(c)

B1cso*: Correct expression with square root **or** correct formula and 5.188 or better
Allow a complete correct method finding $\sum x^2 = \text{awrt } 98720$ and $\sigma = \sqrt{\frac{98715.9...}{184} - \left(\frac{4153.6}{184}\right)^2}$

(d)

B1: Identifying z-value for 10th or 90th percentile (allow awrt (±) 1.28)
or for identifying $[P_{90} =]29.251...$ (awrt 29.3) or $[P_{10} =]15.948...$ (awrt 15.9)
(This may be implied by a correct answer awrt 13.3)
M1: for $2 \times z \times 5.19$ where $1 < z < 2$
or for their $P_{90} - P_{10}$ where $25 < P_{90} < 35$ and $10 < P_{10} < 20$
A1: awrt 13.3

(e)

B1: for one variable identified and a correct supporting reason
B1: for two variables identified and a correct supporting reason for each
Allow any two of the following:

- Wind speed/Beaufort since the data is non-numeric (o.e.). They need not mention Beaufort provided there is a description of the data as non-numeric (Do not allow wind direction/wind gust)
- Rainfall as not symmetric/is skewed/is not bell shaped/lots of 0s /many days with no rain/mean≠mode or median
- Date since each data value appears once/it is uniformly distributed
- Daily mean pressure since it is not symmetric/is skewed/not bell shaped
- Daily mean wind speed since it is not symmetric/is skewed/not bell shaped

Do not allow 'not continuous' or 'discrete' as a supporting reason.
Ignore extraneous non-contradicting statements

Question	Scheme	Marks	AOs

Question	Scheme		Marks	AOs
3(a)	$H_0 : \rho = 0 \qquad H_1 : \rho > 0$		B1	2.5
	Critical value 0.3438		M1	1.1a
	(0.446 > 0.3438) so there is evidence that the product moment correlation coefficient (pmcc) is greater than 0/there is positive correlation		A1	2.2b
			(3)	
(b)	The value is close(r) to 1 **or** there is strong(er) (positive) correlation		B1	2.4
			(1)	
(c)	$\log_{10} y = -1.82 + 0.89(\log_{10} x)$	$y = ax^n \rightarrow$ $\log_{10} y = \log_{10}(ax^n)$	M1	1.1b
	$y = 10^{-1.82+0.89(\log_{10} x)}$	$\log_{10} y = \log_{10} a + \log_{10} x^n$	M1	2.1
	$y = 10^{-1.82} \times 10^{0.89(\log_{10} x)}$ $[= 10^{-1.82} \times 10^{(\log_{10} x)^{0.89}}]$	$\log_{10} y = \log_{10} a + n\log_{10} x$ $[\log_{10} a = -1.82, n = 0.89]$	M1	1.1b
	$y = 0.015x^{0.89}$	$y = 0.015x^{0.89}$	A1A1	1.1b 1.1b
			(5)	
				(9 marks)

Notes

(a)

B1: for both hypotheses correct in terms of ρ

M1: for the critical value: sight of 0.3438 or any cv such that $0.25 < |cv| < 0.45$

A1: a comment suggesting a significant result/ H_0 is rejected on the basis of seeing +0.3438 **and** which mentions "pmcc/correlation/relationship" and "greater than 0/positive" (not just $\rho > 0$)

or an answer in context e.g. 'as "income"(o.e.) increases, "CO_2/emissions"(o.e.) increases'

A **contradictory** statement scores A0 e.g. 'Accept H_0, therefore positive correlation'

(b)

B1: for suitable reason e.g. *r* is close(r) to 1 **or** "strong(er)"/"near perfect" "correlation"

Do not allow 'association'

(c)

For both methods, once an M0 is scored, no further marks can be awarded and condone missing base 10 throughout

Method 1: (working to the model)

M1: Correct substitution for both *c* and *m* (may be implied by 2nd M1 mark)

M1: Making *y* the subject to give an equation in the form $y = 10^{a+b(\log_{10} x)}$ (may be implied by 3rd M1 mark)

M1: Correct multiplication to give an equation in the form $y = 10^a \times 10^{b(\log_{10} x)}$ (this line implies M1M1M1 provided no previous incorrect working seen)

Method 2: (working from the model)

M1: Taking the log of both sides (may be implied by 2nd M1 mark)

M1: Correct use of addition rule (may be implied by 3rd M1 mark)

M1: Correct multiplication of power (this line implies M1M1M1 provided no previous incorrect working seen)

A1: $n = 0.89$ **or** $a =$ awrt 0.015 **or** $y = ax^{0.89}$ **or** $y =$ awrt$0.015x^n$ (**dep on M3**)

A1: $n = 0.89$ **and** $a =$ awrt 0.015 / $y =$ awrt$0.015x^{0.89}$ (**dep on M3**)

do not award the final A1 if answer is given in an incorrect form e.g $y = 0.015 + x^{0.89}$

Question	Scheme	Marks	AOs

Qu 4	Scheme	Marks	AO
(a)	Convenience <u>or</u> opportunity [sampling]	B1 (1)	1.2
(b)	Quota [sampling] e.g. Take 4 people every 10 minutes	B1 B1 (2)	1.1a 1.1b
(c)	Census	B1 (1)	1.2
(d)	[58 – 26 =] **<u>32</u>** (min)	B1 (1)	1.1b
(e)	$\mu = \frac{4133}{95} = 43.505263\ldots$ **awrt <u>43.5</u> (min)** $\sigma_x = \sqrt{\frac{202\,294}{95} - \mu^2} = \sqrt{236.7026...}$ = 15.385… **awrt <u>15.4</u> (min)**	B1 M1 A1 (3)	1.1b 1.1b 1.1b
(f)	There are outliers in the data (or data is skew) which will affect mean and sd Therefore use median and IQR	B1 dB1 (2)	2.4 2.4
(g)	Value of 20, LQ at 26 and outliers will not change <u>or</u> state that median and upper quartile are the values that <u>do</u> change <u>More values now below 40 than above</u> so Q_2 <u>or</u> Q_3 will change and be lower Both Q_2 <u>and</u> Q_3 will be lower	B1 M1 A1 (3)	1.1b 2.1 2.4
		(13 marks)	

	Notes
(b)	1st B1 for quota (sampling) mentioned ("Stratified" or "systematic" or "random" are B0B0) 2nd B1 for a description of how such a system might work, requires suitable strata or categories e.g. time slots, departments, gender, age groups, distance travelled etc Suggestion of randomness is B0
(e)	B1 for a correct mean (awrt 43.5) M1 for a correct expression for the sd (including $\sqrt{\ }$)ft their mean A1 for awrt 15.4 (Allow s = 15.4667… awrt 15.5)
(f)	1st B1 for acknowledging <u>outliers</u> or <u>skewness</u> are a problem for <u>mean and sd</u> "extreme values"/"anomalies" OK May be implied by saying median and IQR not affected by.. We need to see mention of "outliers", "skewness" and the problem so "data is skewed so use median and IQR" is B0 unless mention that they are not affected by extreme values <u>or</u> mean and standard deviation can be "inflated" by the positive skew etc 2nd dB1 dep on 1st B1 for therefore choosing <u>median and IQR</u>
(g)	B1 for identifying 2 of these 3 groups of unchanged values or stating only Q_2 and Q_3 change M1 for <u>explaining</u> that median or UQ should be lower. E.g. the 2 values have moved to below 40 (or 58) and therefore more than 50% below 40 or (more than 75% below 58) <u>or</u> an argument to show that the other 3 values are the same. (o.e.) Allow arrows on box plot provided statement in words about increased % below 40 or 58 etc A1 for stating median <u>and</u> UQ are both lower with clear evidence of M1 scored [If lots of values on 40 then median might not change but, since two values <u>do</u> change then UQ would change. If this meant that 92 became an outlier then we would have a new value for upper whisker and an extra outlier so effectively 3 values are altered. So median changes]