SRCC/PMCC Mark Scheme

Q1.

Question number	Answer	Additional guidance	Mark
(a)	B2 SRCC is sensible because... (one reason from) • e.g. we expect correlation (if price • depends on quality) e.g. it will show if there is correlation / is a relationship e.g. the (three) highest quality mince pies do seem to be the (three) most expensive	B2 for correct conclusion with a sensible reason why. OR B1 for an incomplete answer. e.g. correct reasoning with incorrect or no conclusion, or 'yes' with an attempt at a reason e.g. 'yes, as it is bivariate/ranked data', is B1 only	(2)
(b)	B1 e.g. higher quality mince pies are more expensive, or price does depend on quality, or there is agreement between (quality and price) ranks B1 there is ... • **positive correlation**, or • (0.77) is **close to 1**, or • accept **strong correlation**	1st B1 for a correct interpretation of (positive) correlation. Allow equivalent wording. (condone e.g. 'taste' for 'quality') Note e.g. 'as one increases the other increases' is B0 2nd B1 for formally recognising (positive) correlation.	(2)

Q2.

Question	Answer	Additional guidance	Mark
(a)	B1 The result is outside the range $-1 \leq r \leq 1$	B1 for correct interpretation of the correlation coefficient, allow the result is bigger than 1	(1)
(b) (i) (b) (ii)	B1 Positive B1 The students were in reasonable/good agreement oe	B1 for correct statistical conclusion B1 for correct contextual interpretation of conclusion	(2)

Q3.

Question number	Answer	Additional guidance	Mark
(a)(i)	B1 Positive correlation	B1 for correct statistical conclusion	(1)
(a)(ii)	B1 The judges were in reasonable/good agreement oe	B1 for contextual interpretation of conclusion	(1)
(b)	B1 for answer with reason, e.g. • No AND reference to the change in context (sponge cakes versus flower arranging) • Yes, likely to give similar ranks AND reference to the fact that the judges have similar tastes		(1)

Q4.

Question number	Answer	Additional guidance	Mark
(a)	M1 difference in ranks: 0, 1, −2, −1, −3, 2, 2, 1, 0 M1 $(r_s =) 1 - \dfrac{6 \times 24}{9 \times (9^2 - 1)}$ A1 0.8 B1ft Positive (rank) correlation B1ft Judges were in agreement with the public	M1 for difference in ranks (condone one slip and allow ±). Can be implied by $\Sigma d^2 = 24$ M1 for demonstrating correct use of Spearman's formula A1 cao B1ft for statistical interpretation of their '0.8' B1ft for correct contextual conclusion from their '0.8'	(5)
(b)(i)	B1 **both negative** values identified for the graph B1 Spearman's = −0.9 AND pmcc = −0.7	B1 for recognising graph will give negative correlation B1 cao, for recognising pmcc calculation will be closer to zero when correlation is non-linear	(3)
(b)(ii)	B1 pmcc is less strong correlation as it measures closeness to a **linear** model	B1 for equivalent statistical reasoning that pmcc will be closer to 0 as graph does not suggest a straight line.	

Q5.

Question	Answer	Additional guidance	Mark
(a)	<table><tr><th>Country</th><th>GDP rank</th><th>d</th><th>Σd^2</th></tr><tr><td>Norway</td><td>2</td><td>−1</td><td>1</td></tr><tr><td>Denmark</td><td>4</td><td>−2</td><td>4</td></tr><tr><td>Iceland</td><td>3</td><td>0</td><td>0</td></tr><tr><td>Switzerland</td><td>1</td><td>3</td><td>9</td></tr><tr><td>Finland</td><td>6</td><td>−1</td><td>1</td></tr><tr><td>Netherlands</td><td>5</td><td>1</td><td>1</td></tr><tr><td>Canada</td><td>7</td><td>0</td><td>0</td></tr><tr><td>New Zealand</td><td>8</td><td>0</td><td>0</td></tr></table> B1 GDP rank 2, 4, 3, 1, 6, 5, 7, 8 M1 differences −1, −2, 0, 3, −1, 1, 0, 0 M1 $1-\dfrac{6\times'16'}{8\times(8^2-1)}$ A1 0.81	B1 for GDP rank correct (allow reversed ranks 7, 5, 6, 8, 3, 4, 2, 1) M1 for difference in ranks (condone one slip and allow ±). Can be implied by $\Sigma d^2 = 16$ (or $\Sigma d^2 = 152$ for reversed ranks) M1 for demonstrating correct use of Spearman's formula A1 for awrt 0.81 (allow awrt − 0.81 for reversed ranks)	(4)
(b)	B1 Positive correlation	B1 for correct description of correlation	(2)
	B1 The wealthier the country the happier the country (for the sampled people)	B1 for correct interpretation of the correlation	
(c)	B1 The result is outside the range $-1 \le r \le 1$	B1 for correct interpretation of the correlation coefficient, allow the result is bigger than 1	(1)
(d)	B1 both values show negative correlation B1 −0.8 shows stronger correlation than − 0.5	B1 for each correct comparison	(2)

Q6.

Question number	Answer	Additional guidance	Mark
	M1 Judge's ranks: 2, 1, 3, 5, 6, 4	1st M1 for correct ranks (accept reversed)	(5)
	M1 $d^2 = 0 + 0 + 4 + 1 + 0 + 1 \ (= 6)$	2nd M1 for attempting sum of squared differences of ranks with at least 4 correct	
	M1 $1 - \dfrac{6 \times 6}{6 \times (6^2 - 1)}$	3rd M1 for complete attempt at formula, including '1 − ' (allow their '6')	
	A1 $= 0.828\ldots$	A1 for 0.83 or better	
	A1ft (positive correlation) so agreement between judge and Mayor.	A1ft for correct interpretation of their SRCC. Dependent on complete attempt to use formula **and** r value in range −1 to +1	

Q7.

Question number	Answer	Additional guidance	Mark
(a)	B1 Any reason from: • faster/convenient/easier • (more) accurate/reduces human error	B1 for any suitable advantage of using technology Do not allow cheap(er).	(1)
(b)(i)	B1 Positive (correlation) B1 'As age increases, salary increases'	B1 for positive B1 for correct interpretation of positive correlation (allow converse statements)	(2)
(b)(ii)	B1 Spearman's/0.95 (since they are both positive) as it is closer to 1	B1 for Spearman/0.95 with correct supporting reason Allow e.g. Spearman/0.95 is larger (than 0.77) as supporting reason	(1)
(c)	B1 Figure 1 B1 e.g. 'Figure 1 as pmcc < Spearman (so that means the correlation will be less linear)'	B1 for Figure 1 depB1 (dependent upon 1st B1) either pmcc/0.77 < Spearman/0.95 or for understanding that **Spearman** shows **rank** correlation and **pmcc** shows **linear** correlation)	(2)
(d)	B1 e.g. 'Not appropriate since **bivariate** data is needed for the pmcc'	B1 for not appropriate with supporting reason stating or implying that the data is not bivariate/paired Allow e.g. 'not appropriate since it will only show the relationship between age and salary'	(1)

Q8.

Question	Answer	Additional guidance	Mark
(a)	B1 for a scatter graph showing positive non-linear correlation		(1)
(b)	B1 for 'as the amount of money spent on ski equipment increases, time to complete the ski course decreases'	B0 if more than one box ticked	(1)
(c)	B1 for eg the PMCC value shows correlation, but not causation		(1)

Q9.

Question number	Answer	Additional guidance	Mark
(a)	B2 for Yes / good choice of diagram AND reference to the data being bivariate OR if B2 not earned B1 for Yes / good choice of diagram and reason that does not refer to the type of data OR B1 for referring to the data as bivariate but without commenting on whether the diagram is appropriate or not	B2 for complete assessment of the appropriateness of the diagram with a reason OR if B2 not earned B1 for an incomplete assessment of the appropriateness of the diagram	(4)
(b)	B1 for Lata is not right – (the test results are correlated, but) the English test mark does not cause the Maths test mark oe	B1 for assessment of the given conclusion, including reference to correlation not implying causation	(1)
(c)	B1 for correct interpretation in context e.g. • Correlation for Maths test scores and Science test scores is stronger than the correlation between Maths test score and English test score (or second correlation is stronger) oe • There is a greater association between Maths test score and Science test score than between Maths test score and English test score	B1 for statistical interpretation in context	(1)

Q10.

Question number	Answer	Additional guidance	Mark

(a)	B1 for both negative (correlations)
	B1 for Spearman's rank correlation coefficient
	  closer to − 1/stronger
		B1 must refer to both
		B1 allow 'Spearman's
		  is lower'
		  if first B1 scored
			(2)

(b)	B1 B1 B1 for any three from:
	
	Correlation:
	• Both have positive correlation
	• The correlation between wheat yield) and barley yield is stronger than between wheat yield and oat yield
	• Wheat is a better predictor of barley yield than it is of oat yield
	• Wheat yield and barley yield fit a linear model better than wheat yield and oat yield
	
	Regression Equations:
	
	B1 for any one from:
	• Gradient of regression equation is greater for wheat yield and oats yield than for wheat yield and barley yield
	• It is meaningless to compare y −intercepts as you cannot have negative yields
	B2 for e.g.
	  Each additional t/ha for wheat would suggest an additional 1.52t/ha
	
	  for oats / Each additional t/ha for wheat would suggest an additional
	
	  1.24t/ha for barley
		B1 for each correct comparison (maximum 3) based on the correlation
		
		B1 for correct comment on either gradient or y-intercept
		
		B2 for contextual interpretation of gradient of regression equation
		
		(B1 for incomplete interpretation e.g. 'As wheat yield increases, both oat yield and barley yield increase')
			(6)

(c)(i)	M1 $1.24 \times \frac{75}{28} - 0.30$ or $1.52 \times \frac{75}{28} - 1.05$ A1 for demonstrating both sides of the equation give the same value. OR M1 $-0.30 + 1.05 = 1.52x - 1.24x$ A1 $\frac{75}{28}$	M1 for substituting into one half of the equation given A1 for demonstrating both sides of the equation give the same value. OR M1 for a method to solve the equation formed A1 given answer from correct working	(2)
(c)(ii)	B2 e.g. if there is less than 2.67.. t/ha of wheat produced then plant barley rather than oats to get the larger yield (B1 e.g. if there is 2.67 t/ha of wheat then barley and oats would be expected to give the same yield)	B2 for contextual interpretation of the point of intersection of the regression equations (B1 for use of line of line of best fit with 2.67...)	(2)
(c)(iii)	B1 For any one from • the data collected related to farms in Australia / only fields in sample • data does not take into account different growing conditions/weather • lower correlation for oat yield implies less confidence in the prediction.	B1 for a correct limitation of the data	(1)

Q11.

Question number	Answer	Additional guidance	Mark
(a)	<table><tr><th>Time ranks</th><th>d (difference in ranks)</th></tr><tr><td>5</td><td>−4</td></tr><tr><td>2</td><td>0</td></tr><tr><td>7</td><td>−4</td></tr><tr><td>(1)</td><td>3</td></tr><tr><td>8</td><td>−3</td></tr><tr><td>6</td><td>0</td></tr><tr><td>4</td><td>3</td></tr><tr><td>3</td><td>5</td></tr></table> M1 M1 $\sum d^2 = 84$ A1 $r_s = 1 - \frac{6 \times 84}{8(8^2-1)} = 0$ B1ft No correlation, so… depB1ft Amelia's hypothesis is not supported	M1 at least 5 correct time ranks (may be implied by 2nd M1) Allow if one rank misplaced but then subsequent ranks in correct order. M1 attempt at calculating sum of d^2 for *their* ranks A1 $r_s = 0$ B1ft for no correlation ft their r_s provided $-1 \le r_s \le 1$ depB1ft not supported (dep on at least 1 previous M mark being scored and an attempt at identification of correlation) allow follow through their value of r_s	(5)
(b)	B1 e.g. 'collect more data', 'repeat 100 metre race more than once'	B1 for a suitable reason to improve the reliability of her results Condone 'collect primary data'	(1)