Question 1:

Tessa owns a small clothes shop in a seaside town. She records the weekly sales figures, £ w, and the average weekly temperature, t°C, for 8 weeks during the summer. The product moment correlation coefficient for these data is -0.915.

(a) Stating your hypotheses clearly and using a 5% level of significance, test whether or not the correlation between sales figures and average weekly temperature is negative.

[3]

(b) Suggest a possible reason for this correlation.

[1]

Tessa suggests that a linear regression model could be used to model these data.

(c) State, giving a reason, whether or not the correlation coefficient is consistent with Tessa's suggestion.

[1]

(d) State, giving a reason, which variable would be the explanatory variable.

[1]

Tessa calculated the linear regression equation as $w = 10755 - 171t$.

(e) Give an interpretation of the gradient of this regression equation.

[1]

Question 2:

A meteorologist believes that there is a relationship between the daily mean windspeed, w kn, and the daily mean temperature, t°C. A random sample of 9 consecutive days is taken from past records from a town in UK in July and the relevant data is given in the table below.

t	13.3	16.2	15.7	16.6	16.3	16.4	19.3	17.1	13.2
w	7	11	8	11	13	8	15	10	11

(a) Explain why a linear regression model based on these data is unreliable on a day when the mean temperature is 24°C.

[1]

(b) State what is measured by the product moment correlation coefficient.

[1]

(c) Stating your hypotheses clearly, test, at the 5% significance level, whether or not the product moment correlation coefficient for the population is greater than zero.

[3]

Using the same 9 days a location from the large data set gave $\bar{t} = 27.2$ and $\bar{w} = 3.5$.

(d) Using your knowledge from the large data set, suggest, giving a reason, the location that gave rise to these statistics.

[1]

Question 3:

Barbara is investigating the relationship between average income (GDP per capita), x US dollars, and average annual carbon dioxide (CO_2) emissions, y tonnes, for different countries.

She takes a random sample of 24 countries and finds the product moment correlation coefficient between average annual CO_2, emissions and average income to be 0.446.

(a) Stating your hypotheses clearly, test, at the 5% level of significance, whether or not the product moment correlation coefficient for all countries is greater than zero.

[3]

Barbara believes that a non-linear model would be a better fit to the data. She codes the data using the coding $m = \log_{10} x$ and $c = \log_{10} y$ and obtains the model $c = -1.82 + 0.89m$. The product moment correlation coefficient between c and m is found to be 0.882.

(b) Explain how this value supports Barbara's belief.

[1]

(c) Show that the relationship between y and x can be written in the form $y = ax^n$ where a and n are constants to be found.

[5]

Question 4:

Kathleen is exploring the large data set. She wants to know if there is a correlation between temperatures in different locations.

Here are the daily mean temperatures for the first 10 days of May 2015 for Beijing and for Perth.

Beijing	17.5	20.0	19.2	18.5	21.1	17.1	18.8	18.0	13.0	9.7
Perth	15.8	16.4	16.1	9.7	12.0	13.8	14.0	15.2	15.0	16.5

(a) Calculate the product moment correlation coefficient for the daily mean temperatures for Beijing and for Perth for the first 10 days of May 2015.

[1]

Kathleen selects a random sample of 30 days from 2015. She calculates the product moment correlation coefficient for daily mean temperatures for these 30 days for Heathrow and Hurn. The product moment correlation coefficient for these data is 0.8198.

(b) Test, using a 5% level of significance, whether or not the correlation between daily mean temperatures for Heathrow and for Hurn is positive.

[3]

For these 30 days she also calculates the product moment correlation coefficient for daily mean temperatures for Heathrow and Leuchars. The product moment correlation coefficient for these data is 0.5612.

(c) Compare the product moment correlation coefficients for daily mean temperatures for Heathrow and Hurn and for Heathrow and Leuchars. You should:

- comment on the strength of any relationships
- give an interpretation of these correlation coefficients
- suggest a reason for the difference in the values of the correlation coefficients

[3]

Question 5:

Jon is using the large data set to carry out investigations into the weather. He randomly selects 10 days from the large data set for Leuchars in 2015. He calculates the product moment correlation coefficient between daily total sunshine, in hours, and daily total rainfall for this sample.

His value is $r = -0.7414$.

(a) Using Jon's data test, at the 5% level of significance, whether or not there is evidence of a negative correlation between daily total sunshine and daily total rainfall at Leuchars in 2015.

[3]

(b) Using your knowledge of the large data set, state the units used for the daily total rainfall.

[1]

Jon wishes to use the daily total sunshine to try to predict the daily total rainfall on a particular day.

(c) State which would be the explanatory variable in this case.

[1]

(d) Explain why Jon's product moment correlation coefficient does not necessarily justify predicting the daily total rainfall from the daily total sunshine.

[1]

Product moment coefficient					
		Level			**Sample**
0.10	**0.05**	**0.025**	**0.01**	**0.005**	**Level**
0.8000	0.9000	0.9500	0.9800	0.9900	4
0.6870	0.8054	0.8783	0.9343	0.9587	5
0.6084	0.7293	0.8114	0.8822	0.9172	6
0.5509	0.6694	0.7545	0.8329	0.8745	7
0.5067	0.6215	0.7067	0.7887	0.8343	8
0.4716	0.5822	0.6664	0.7498	0.7977	9
0.4428	0.5494	0.6319	0.7155	0.7646	10
0.4187	0.5214	0.6021	0.6851	0.7348	11
0.3981	0.4973	0.5760	0.6581	0.7079	12
0.3802	0.4762	0.5529	0.6339	0.6835	13
0.3646	0.4575	0.5324	0.6120	0.6614	14
0.3507	0.4409	0.5140	0.5923	0.6411	15
0.3383	0.4259	0.4973	0.5742	0.6226	16
0.3271	0.4124	0.4821	0.5577	0.6055	17
0.3170	0.4000	0.4683	0.5425	0.5897	18
0.3077	0.3887	0.4555	0.5285	0.5751	19
0.2992	0.3783	0.4438	0.5155	0.5614	20
0.2914	0.3687	0.4329	0.5034	0.5487	21
0.2841	0.3598	0.4227	0.4921	0.5368	22
0.2774	0.3515	0.4133	0.4815	0.5256	23
0.2711	0.3438	0.4044	0.4716	0.5151	24
0.2653	0.3365	0.3961	0.4622	0.5052	25
0.2598	0.3297	0.3882	0.4534	0.4958	26
0.2546	0.3233	0.3809	0.4451	0.4869	27
0.2497	0.3172	0.3739	0.4372	0.4785	28
0.2451	0.3115	0.3673	0.4297	0.4705	29
0.2407	0.3061	0.3610	0.4226	0.4629	30

Numerical Answers:

(1) (a) Sufficient evidence to suggest negative correlation between w and t

(b) As temperature increases people spend more time on the beach and less time shopping

(c) r is close to -1, so it is consistent

(d) t is explanatory variable, since sales depends on temperature

(e) Sales decreases by £ 171 for every increase in temperature by 1°C

(2) (a) Extrapolation required, hence unreliable

(b) Linear association between w and t

(c) Sufficient evidence to suggest positive correlation between w and t

(d) Higher $\bar{t}$, so overseas; lower $\bar{w}$, so not close to sea; hence Beijing

(3) (a) Sufficient evidence to suggest positive correlation between x and y

(b) This value is closer to 1, so it suggests a stronger correlation

(c) $y = 0.015x^{0.89}$

(4) (a) $r = -0.362$

(b) Sufficient evidence to suggest positive correlation between the mean air temperatures for June 2015 for Heathrow and for Hurn

(c)
- Stronger positive correlation between daily mean air temperatures in June 2015 for Heathrow and Hurn than for Heathrow and Leuchars
- As temperatures in Heathrow increase so do temperatures in Hurn and Leuchars, but association is closer between Heathrow and Hurn than between Heathrow and Leuchars
- Heathrow is closer to Hurn than to Leuchars so weather tends to be more similar

(5) (a) Sufficient evidence to suggest negative correlation between daily total sunshine and daily total rainfall at Leuchars in 2015

(b) Millimetres

(c) Daily total sunshine

(d) Sample is small / Correlation does not imply causation